2017 THE ALMOND CONFERENCE

TECHNOLOGY IN THE FOOD SAFETY WORLD: TOOLS SUCH AS WHO GENOME SEQUENCING – FRIEND OR FOE?

Room 314 | December 5 2017

california almonds®
Almond Board of California

growing ADVANTAGE®
The Almond Conference

CEUs – New Process

Certified Crop Advisor (CCA)

- Sign in and out of each session you attend.
- Pickup verification sheet at conclusion of each session.
- *Repeat this process for each session, and each day you wish to receive credits.*

Pest Control Advisor (PCA), Qualified Applicator (QA), Private Applicator (PA)

- Pickup scantron at the start of the day at first session you attend; complete form.
- Sign in and out of each session you attend.
- Pickup verification sheet at conclusion of each session.
- Turn in your scantron at the end of the day at the last session you attend.

Sign in sheets and verification sheets are located at the back of each session room.

AGENDA

- **Tim Birmingham**, Almond Board of California, moderator
- **Jesse Miller**, NSF International
- **Maria Hoffmann**, FDA Center for Food Safety and applied Nutrition

Next Generation Sequencing – The Technology and its Applications – Friend or Foe?

Jesse D. Miller, Ph.D.
Director
Applied Research Center
NSF AuthenTechnologies

Agenda

Next Generation Sequencing

Methods

Applications and Examples

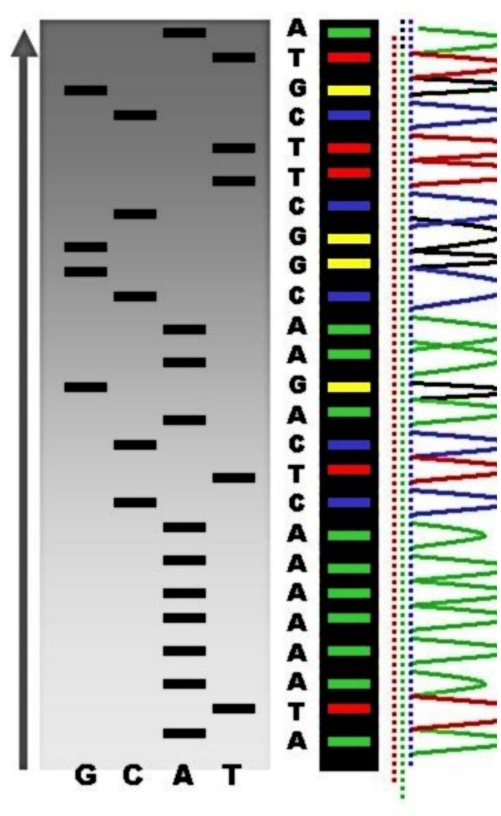
Next Generation Sequencing

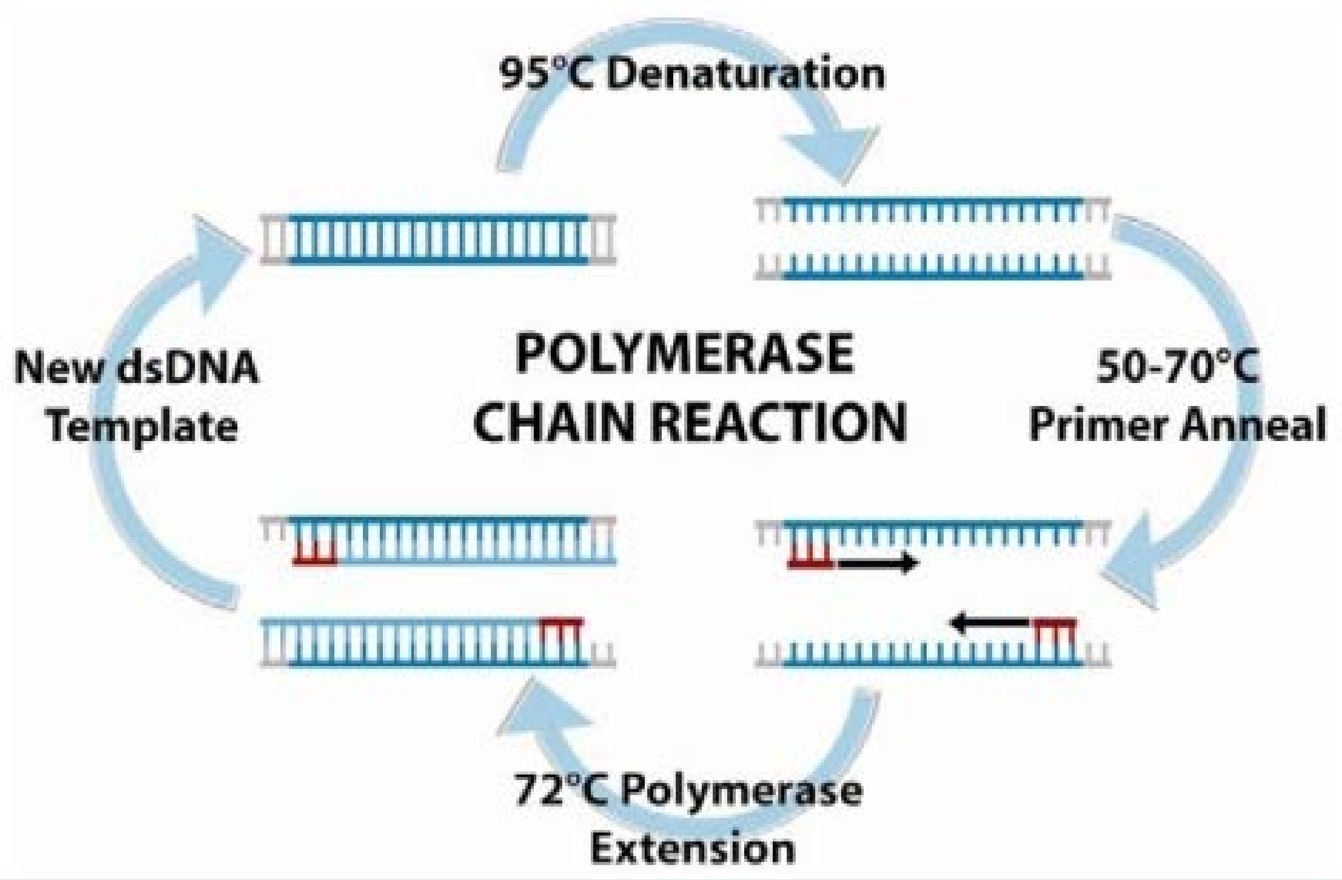
Process of extracting genetic material and reading the “code”.

Lets go Back in Time…………..1952.

How do we Analyze DNA? Sequencing Background

- RNA sequencing was first to be developed (different methods)





What is Next-Gen Sequencing?

- Term used for sequencing that has a higher throughput than traditional Sanger sequencing
- Now Encompasses many platforms – ThermoFisher, Illumina, Pacific Biosciences, Oxford Nanopore
- **Can be Whole Genome Sequencing, 16S rRNA Metagenomics, Shotgun Metagenomics, Targeted Gene Sequencing, RNA-SEQ**
- 1st Gen – Sanger, ABI (3130xl)
- 2nd Gen – 454, Illumina (Solexa) and ThermoFisher (Massively Paralell Sequencers – Short Read)
- 3rd Gen – Pacific Biosciences, Oxford (Long Read Sequencers)

BASES TO BYTES

How is NGS Different than Traditional Sequencing?

- Sequencing done on flowcells/chips now. No 2D gels or capillaries required
- MUCH more data generated (Terabases now, kilobases then)
- Sophisticated BioInformatic programs exist to parse out the data - In some instances, can sequence a sample for around $40
  - MUCH cheaper than historical
- Open source data sharing for massive datasets
- Cloud computing capability

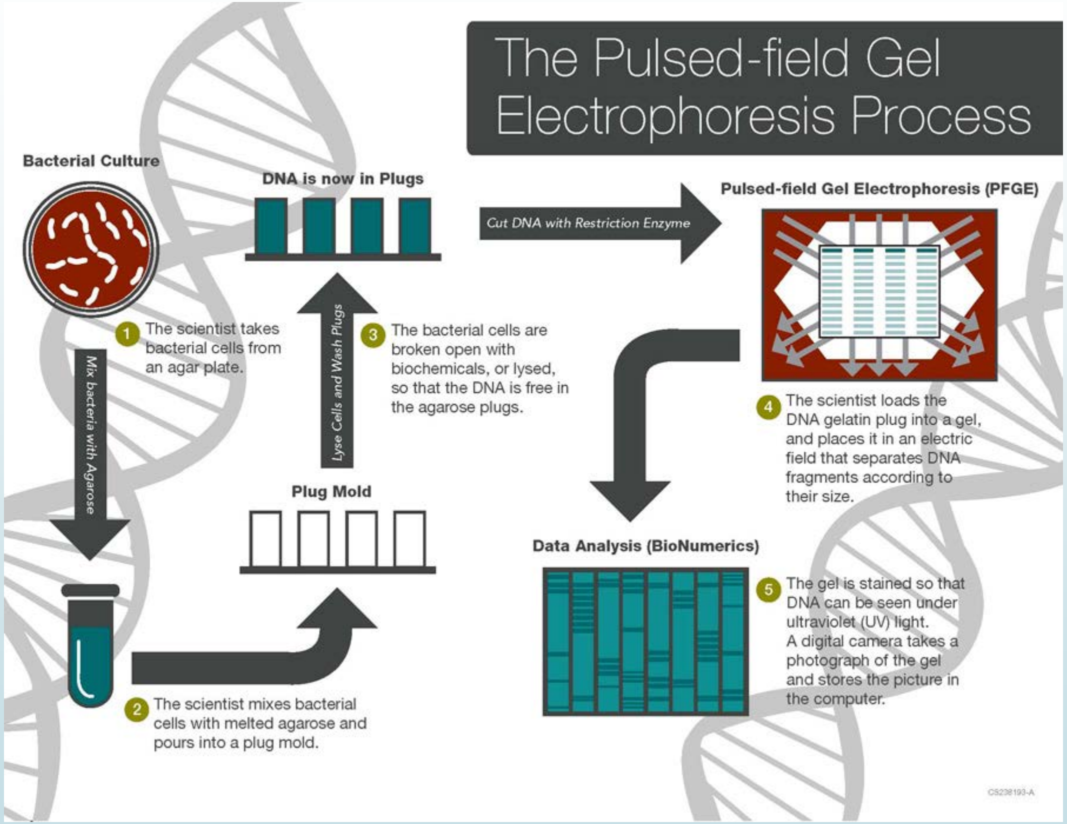
MORE DATA
CHEAPER PER BASE
INTERNET MAKES COMMS AND ANALYSIS EASY

Methods

Choose the Right "Fit for Purpose" Tool for the Job

Pulse-Field Gel Electrophoresis (PFGE)

- “Gold Standard” of bacterial DNA fingerprinting
- Restriction enzymes cut bacterial DNA in specific locations
- Multi-directional gel electrophoresis produces unique pattern based on the fragment sizes
- Allows Comparisons between organisms for ID
- Not quantitative
- $100-260



The Pulsed-field Gel Electrophoresis Process

① The scientist takes bacterial cells from an agar plate.

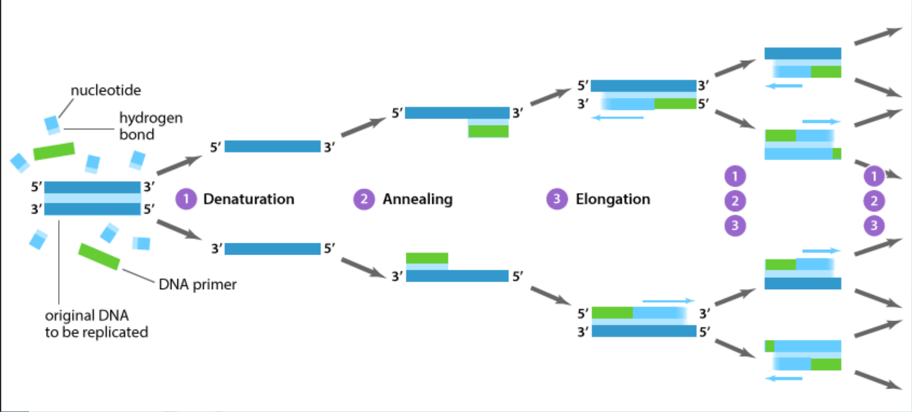
② The scientist mixes bacterial cells with melted agarose and pours into a plug mold.

③ The bacterial cells are broken open with biochemicals, or lysed, so that the DNA is free in the agarose plugs.

④ The scientist loads the DNA gelatin plug into a gel, and places it in an electric field that separates DNA fragments according to their size.

⑤ The gel is stained so that DNA can be seen under ultraviolet (UV) light. A digital camera takes a photograph of the gel and stores the picture in the computer.

Polymerase Chain Reaction



Semi Quantitative (With a standard curve)

Targets a region of genome for amplification

- Positive reaction = gene is present

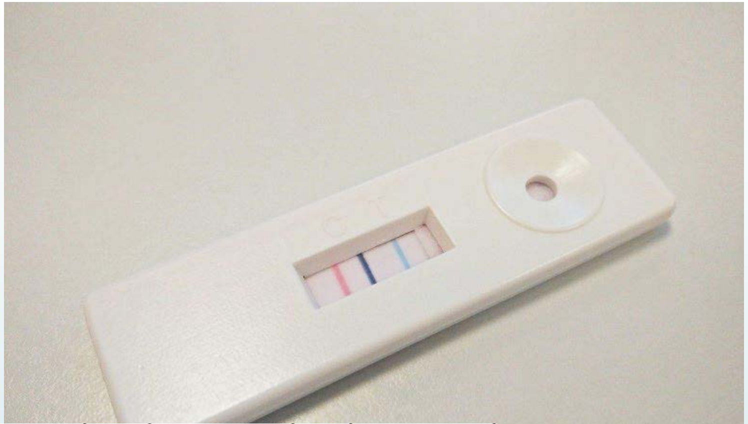
Can detect several target genes at once (Multiplex)

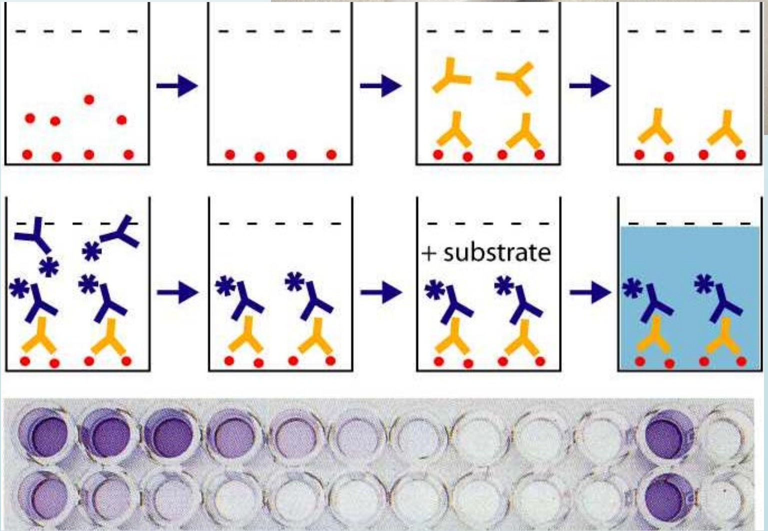
Cheap! $5-10/reaction

Several hours to run

Immunological Methods

- ELISA
- Lateral Flow
- Cheap ($5-10)
- Fast
- Yes/No answers
  - ELISA can be quantitative





What is Whole Genome Sequencing?

- Whole Genome Sequencing is the term used for extraction of DNA from an organism and the subsequent mapping of its genome.

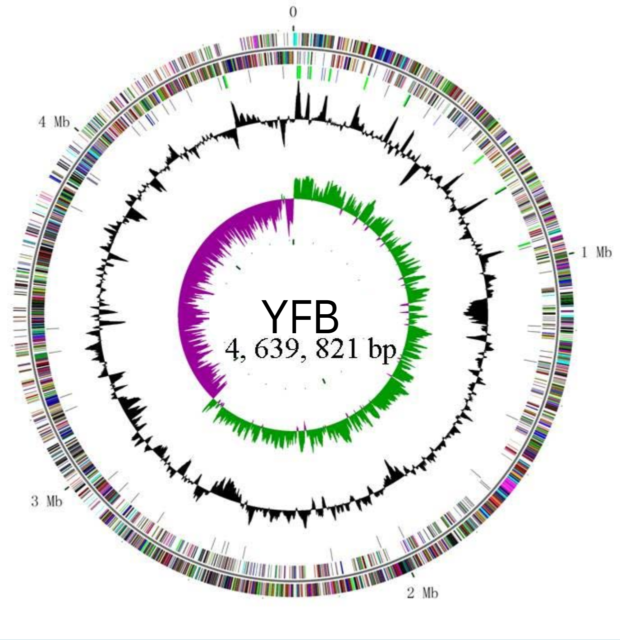
- The genetic code (AGCT) is read on an instrument and written into a digital file.

- That digital file can be assembled (like a linear puzzle) to determine the order of the code in the organism.

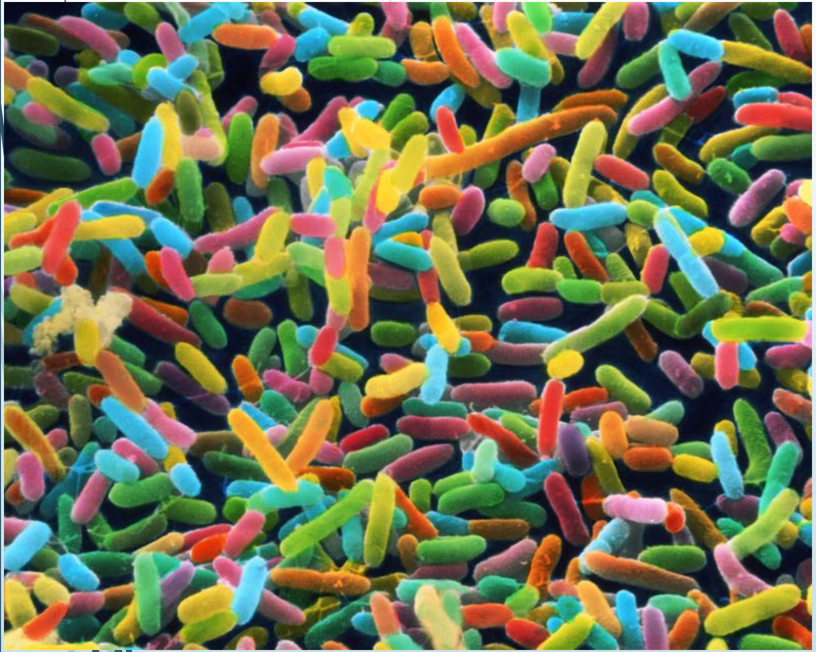
- Once you have the ordered code, you can analyze the data and make comparisons and data-driven decisions about the organism.

- Not quantitative

- $50/sequence – Up to $500 for assembly/closure (Bacteria)



What is 16S/Shotgun Metagenomics?



- 16S Sequencing on Next-Gen platforms follows a similar workflow, except that it targets a specific gene (16S Ribosomal RNA) used to identify bacteria
- Metagenomics applies this concept to mixed consortia, resulting in a profile of bacteria abundance (e.g., your microbiome)

1. Sequence DNA
2. Alignment to reference databases. Classifying unknown bacteria in to taxonomic groups
3. Visualize in phylogenetic trees, pie charts, or other analyses based on question to be asked

POPULATIONS

NGS Benefits Over Other Methods Moving Forward?

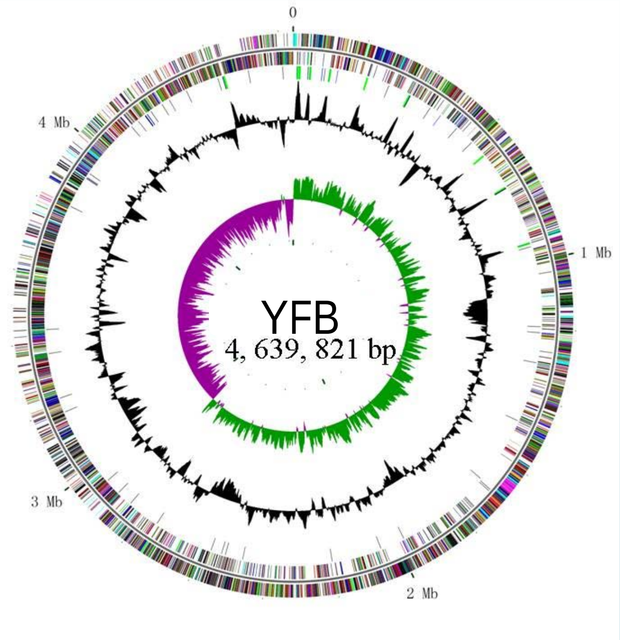
- NGS differentiation (resolution) is unmatched
- With high-throughput efficiencies, NGS is cheaper and faster
- NGS enables much more in-depth data analysis, such as functional genes and heredity
- Cost will continue to decrease
- Database availability and power will continue to increase
- Global adoption and data sharing will increase value

Applications and Examples

Whole Genome Sequencing

- What can I use it for?
  - Epidemiology
  - Resistance
  - Strain level ID
  - Authenticity

- MUCH deeper look into genome than Pulse Field Gel Electrophoresis or RFLP
  - Looking at every base, not just where enzymes cut

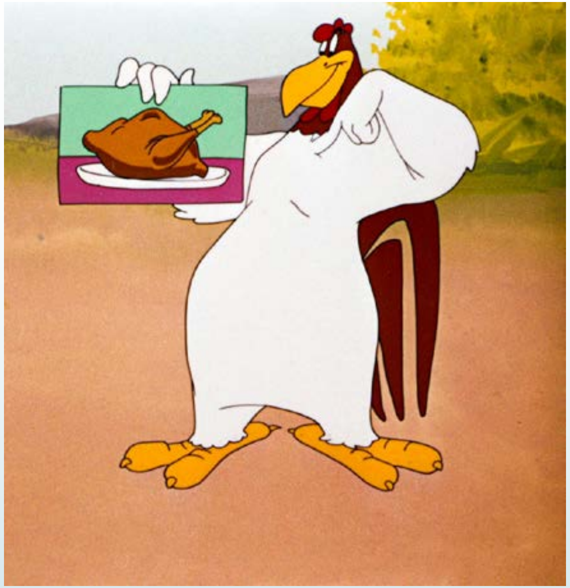


```
>Sequence_2
CCCTAAACCCTAAACCCTAAACCCTAAACCTCTGAATCCTTAATCCCTAAATCCCTAAAT
CTTTAAATCCTACATCCATGAATCCCTAAATACCTAATTCCCTAAACCCGAAACCGGTTT
CTCTGGTTGAAAATCATTGTGTATATAATGATAATTTTATCGTTTTTATGTAATTGCTTA
TTGTTGTGTGTAGATTTTTTAAAAATATCATTTGAGGTCAATACAAATCCTATTTCTTGT
GGTTTTCTTTCCTTCACTTAGCTATGGATGGTTTATCTTCATTTGTTATATTGGATACAA
GCTTTGCTACGATCTACATTTGGGAATGTGAGTCTCTTATTGTAACCTTAGGGTTGGTTT
ATCTCAAGAATCTTATTAATTGTTTGGACTGTTTATGTTTGGACATTTATTGTCATTCTT
```

Strain Level ID

- Is there value in knowing who your resident strains are?
  - Can you eradicate them more easily?
  - Can you modify processes and cleaning regimens?
- Proactivity?
  - Value in transparency and ownership?
    - Working toward a positive solution
- Third Party sequencing
  - Metadata housed by third party

Speciation of *Campylobacter*



Commensal bacteria on Chicken and other fowl

Interventions can knock down numbers, but hard to completely eradicate

Three strains under scrutiny

- Jejuni
- Coli
- Lari

Food Pathogen ID – Off the Shelf

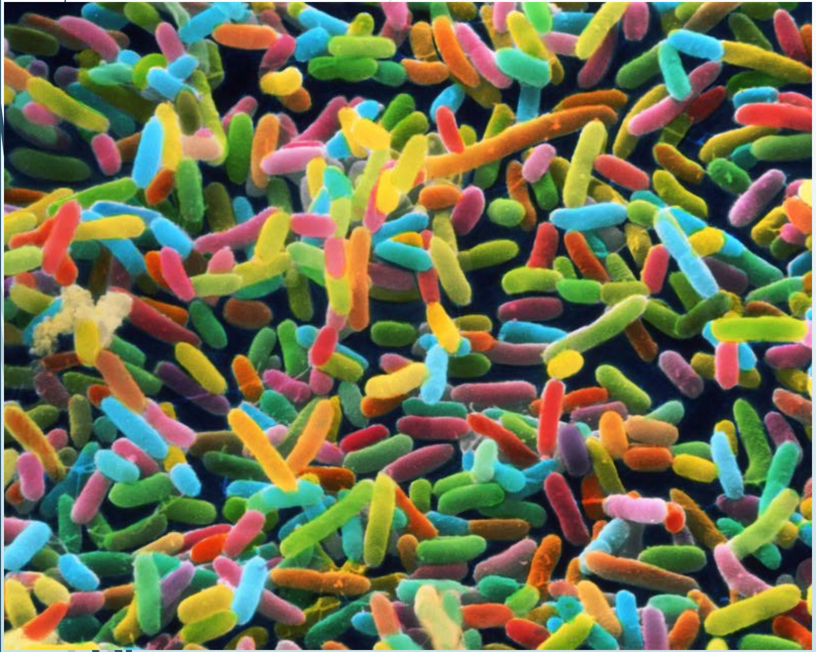
Isolate germs from off the shelf meats

Assay for pathogens

Whole Genome Sequence for species

- Common trends?
  - Food type
  - Geography
  - Intervention method
  - Preservation method

What is 16S/Shotgun Metagenomics?



- 16S Sequencing on Next-Gen platforms follows a similar workflow, except that it targets a specific gene (16S Ribosomal RNA) used to identify bacteria
- Metagenomics applies this concept to mixed consortia, resulting in a profile of bacteria abundance (e.g., your microbiome)

1. Sequence DNA
2. Alignment to reference databases. Classifying unknown bacteria in to taxonomic groups
3. Visualize in phylogenetic trees, pie charts, or other analyses based on question to be asked

POPULATIONS

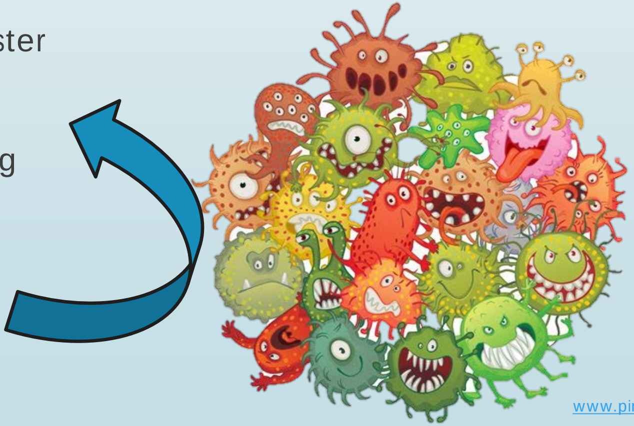
Irrigation Water Microbiome

Looking at changes in water microbiome when *E. coli* present

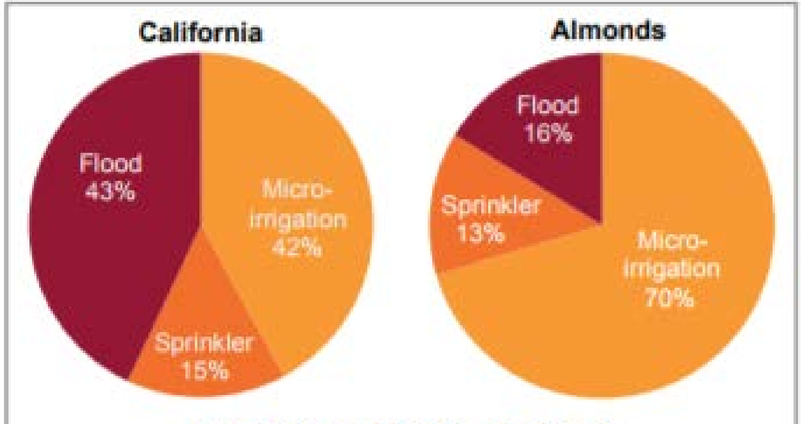
- Searching for markers of contamination

Almond Harvest

- Shaking trees to release fruit
- Drying for a few days
- Harvester
- Hulling
- Shelling



www.pinterest.com



Irrigation Method Used

Source: CA DWR. California Water Plan Update 2013.

Hospital Microbiome

Looking at environment and patient colonization

- Searching for correlations to understand flora

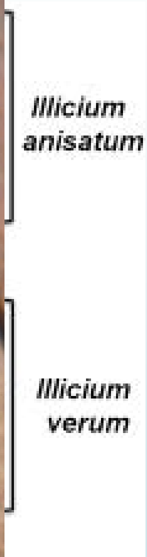
Proactive Treatment?

- Our microbiome protects us – keeps the bad players out
- Antibiotics kill our normal flora
- Probiotic treatment can prevent undesirable bugs from taking hold
- Understanding what is out there allows decision making with more cards in your deck

www.pinterest.com

Epidemiology

- Track and trace your strains
- Look for Single Nucleotide Polymorphisms
  - These changes happen in an organism over time
  - Differentiates one bacteria from another
- GenomeTrakr Application
  - A Database of organisms that can be mined to determine source and traceback
    - Publicly available!


Illicium anisatum
Illicium verum

Next Generation Sequencing – Friend!

Sequencing is the Future of Food Safety and Microbial Science

Not Scary! Just a way to get detailed information about the organism you are analyzing

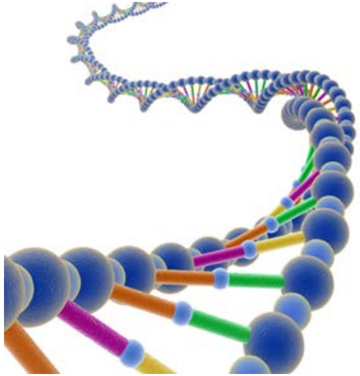
- Can use genomic info to understand
  - Resistances
  - Phylogeny
  - Pathogenesis
  - Epidemiology
- Authenticity
- Make Data Driven Decisions
- Sensitivity – Better Decisions, Faster. Saves lives!

Next Generation Sequencing – The Technology and its Applications – Friend!

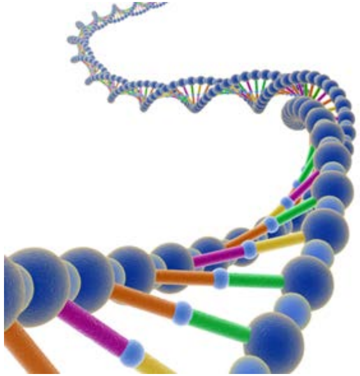
Jesse D. Miller, Ph.D.
JDMILLER@NSF.ORG
734.707.5413

FDA U.S. FOOD & DRUG ADMINISTRATION

Technology in the Food Safety World: Whole Genome Sequencing—friend or foe?



The Almond Conference
Sacramento California
December 5, 2017



Maria Hoffmann, Ph.D.
Genomics Research Microbiologist

Office of Regulatory Science

Tracking contamination down…and FAST!

Global Point Source

Import Lines

Ecologic Reservoirs

Finished Product

Processing Facility

Farm

SAVES LIVES

Some perspective on the food supply

- **Tracking and Tracing of food pathogens**
    - Almost 200,000 registered food facilities (2/14)
        - 81,574 Domestic and 115,753 Foreign
    - More than 300 ports of entry
    - More than 130,000 importers and more than 11 million import lines/year
    - In the US there are more than 2 million farms

The Complex Etiology of Foods

Salad

Shrimp – India
Cilantro – Mexico
Romaine – Salinas, CA
Cheddar – Wisconsin
Carrots – Idaho
Gruyere – Switzerland
Pecans – Georgia
Sprouts – Chicago
Red Cabbage - NY

Sushi

Shrimp – Indonesia
Imitation Crab – Alaska
Tuna Scrape – India
Fish Roe – Seychelles
Salmon – Puget Sound
Soy Sauce – China
Rice – Thailand
Seaweed Wrap – CA
Avocado – Mexico
Cucumber – Maryland
Wasabi – Japan
Pepper – Vietnam

Fruit platter

Watermelon – Delaware
Blackberries – Guatemala
Blueberries – New Jersey
Pineapple – Guam
Grapes – California
Kiwi – New Zealand
Apples – New York
Pears – Oregon
Cantaloupe – Costa Rica
Honeydew – Arizona
Papaya – Mexico
Banana – Costa Rica

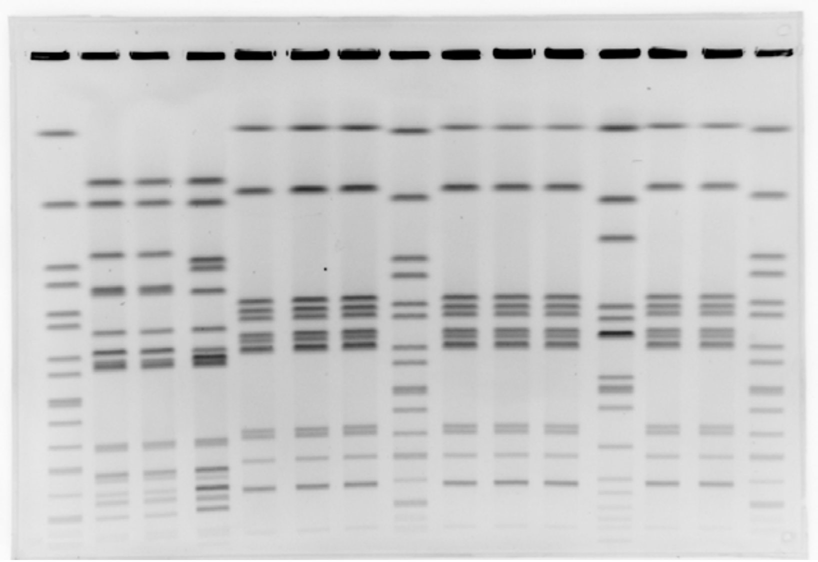
Gold standard method for pathogen identification

PulseNet USA

The National Molecular Subtyping Network for Foodborne Disease Surveillance

PulseNet, est. 1996
http://www.cdc.gov/pulsenet/

PFGE: banding patterns determine discrimination within serovar.



PFGE v/s WGS

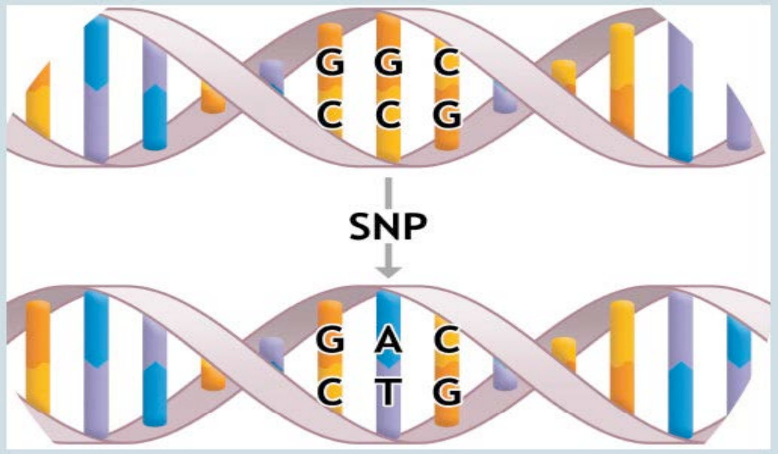
- **WGS is high resolution**

  3-5 million data points are collected for each isolate

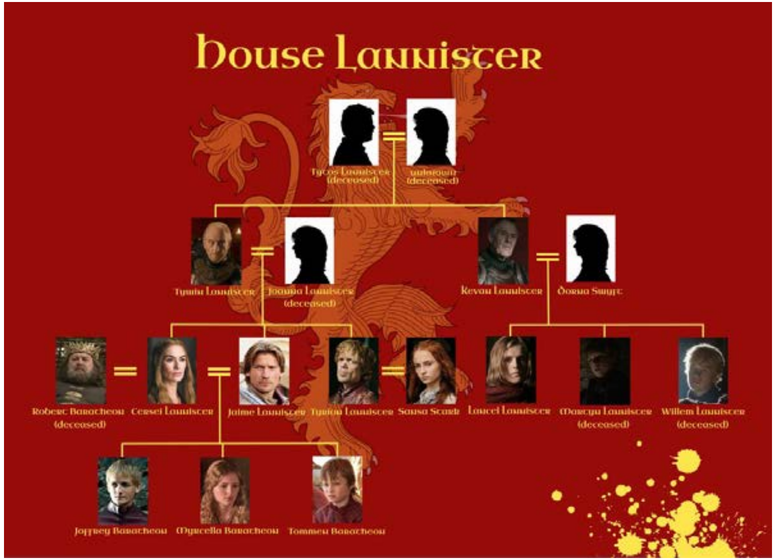
- **WGS analyses are statistically robust**

  Unlike PFGE patterns, WGS data can be analyzed in its evolutionary context.

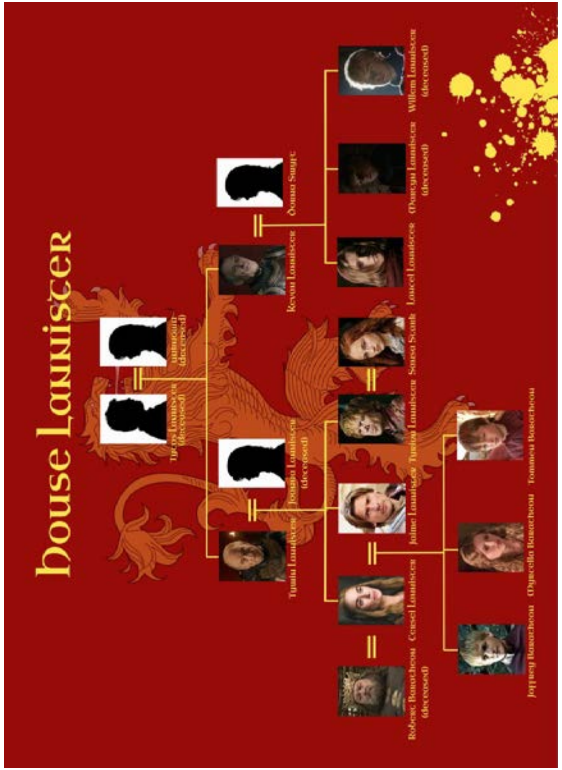
  Accurate and stable genetic changes within pathogen genomes enable us to pin point specific common sources of outbreak strains (farms, processing plants, food types, and geographic regions)



Pedigree vs Phylogeny

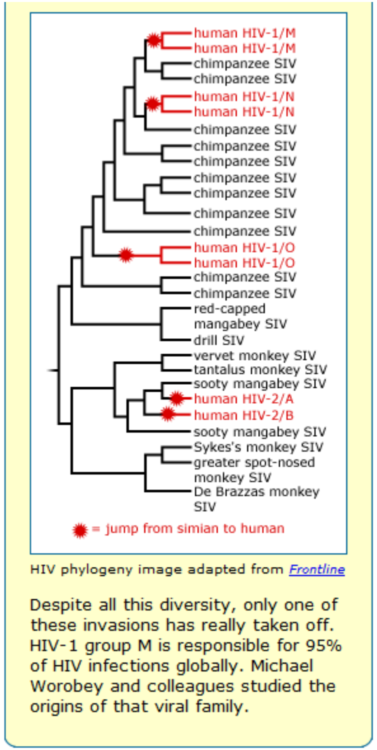


→



DNA based pathogen surviellance not new

- **Flu**: 1990s – flu vaccines predicted from phylogenetic trees
- **HIV**: 1990s – early tracking of HIV transmission using phylogenetics



HIV phylogeny image adapted from *Frontline*

Despite all this diversity, only one of these invasions has really taken off. HIV-1 group M is responsible for 95% of HIV infections globally. Michael Worobey and colleagues studied the origins of that viral family.

http://evolution.berkeley.edu/evolibrary/news/081101_hivorigins

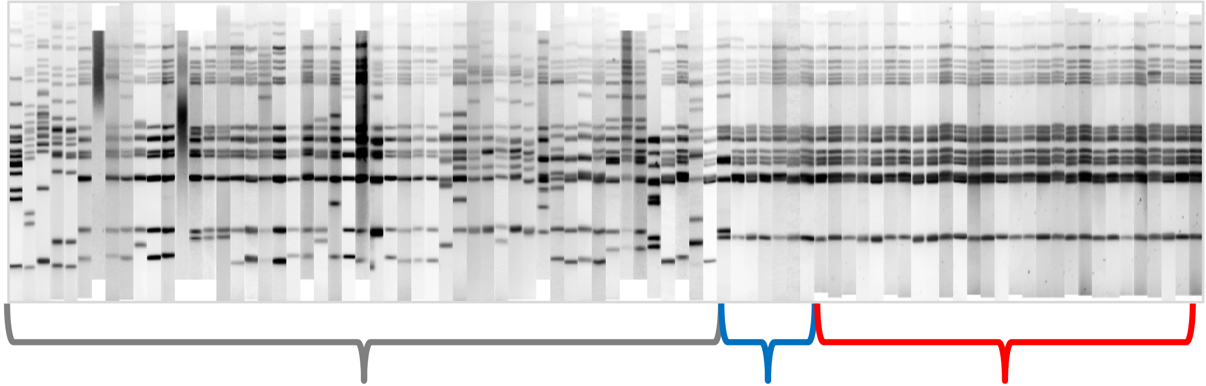
Salmonella enterica serovar Bareilly

- CDC investigated a multistate (29 states) outbreak
- 410 confirmed cases between January 1st and July 7th, 2012
- Among the 326 case patient, 55 (17%) had been hospitalized
- Yellowfin tuna was implicated as source of this outbreak
- This product had been imported from an Indian corporation and was used to make spicy tuna sushi for restaurants and grocery stores
- At this time no reference genome

  was available at NCBI

PFGE identical in red

NGS distinguishes geographical structure among closely related *Salmonella* Bareilly strains



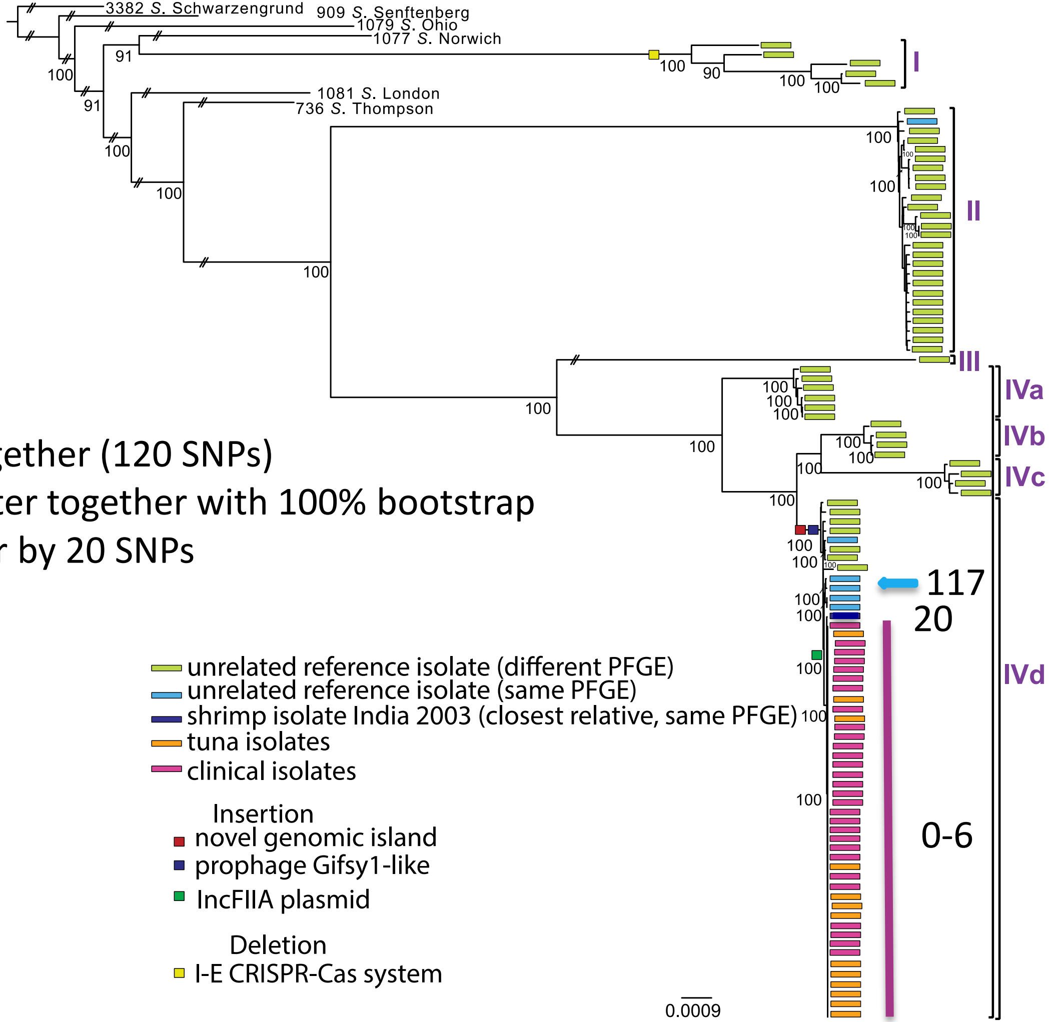
Different PFGE than the outbreak pattern

Same PFGE but not part of the outbreak

Outbreak Isolates

MD isolates – in green

NY isolates – in purple



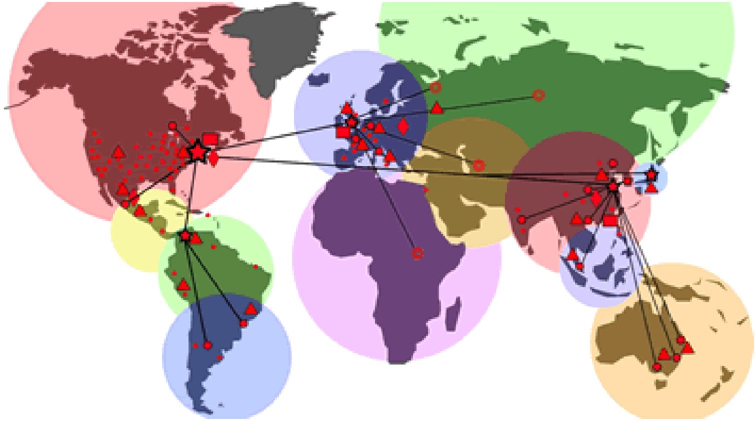
- Same PFGE cluster together (120 SNPs)
- outbreak isolates cluster together with 100% bootstrap
- Closest neighbor differ by 20 SNPs

2-part paradigm shift

1. **Whole genome sequencing**
   - High resolution data
   - Harness established field of evolutionary theory for analyses
1. **Open data**
   - Raw genome sequences made available to the public 1-2 days after collection
   - Data made public *before* FDA analyses are preformed

Why Develop a WGS Based Network?

- **Tracking and Tracing of food pathogens**
  - Insufficient resolution of current tools
    -matching clinical to environmental
  - Faster identification of the food involved in the outbreak
  - Limited number of investigators vs. facilities and import lines
  - Global travel
  - Global food supply



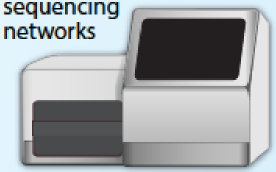
Basic Data Flow for Global WGS Public Access Databases

DATA ACQUISITION

Sequence and upload genomic and geographic data

GenomeTrakr

Other distributed sequencing networks

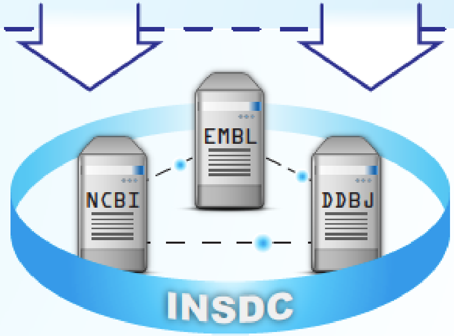


DATA ASSEMBLY, ANALYSIS, AND STORAGE

International Nucleotide Sequence Database Collaboration (INSDC)

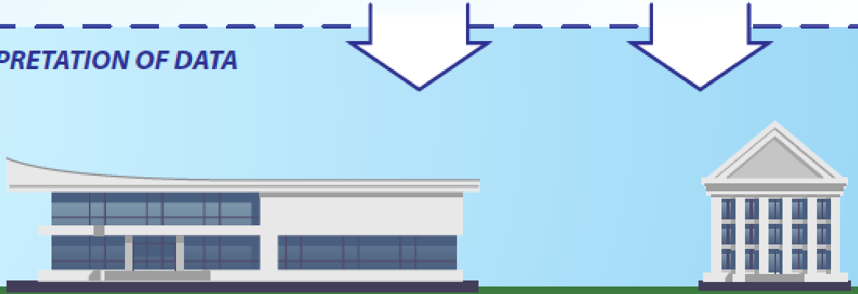
Shared Public Access Databases

- *NCBI – National Center for Biotechnology Information*
- *EMBL – European Molecular Biology Laboratory*
- *DDBJ – DNA Databank of Japan*



PUBLIC HEALTH APPLICATION AND INTERPRETATION OF DATA

- *Find clinical links*
- *Identify clusters*
- *Conduct traceback*
- *Develop rapid methods*
- *Develop culture independent tests*
- *Develop new analytical software*



11/2014

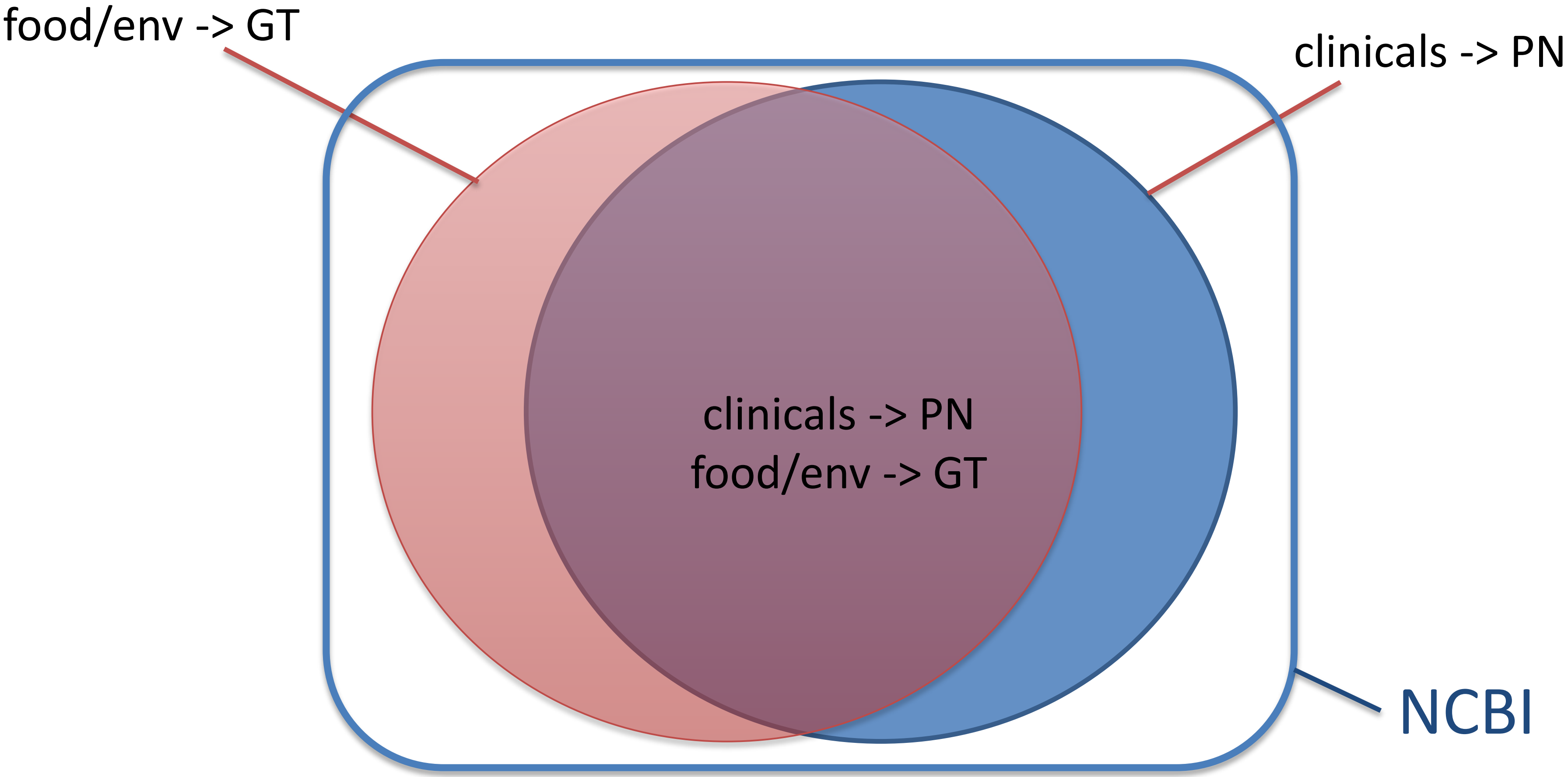
State, Local, Federal, and Foreign Public Health Agencies

Academia/Industry

FDA’s GenomeTrakr

- Distributed network of labs to use whole genome sequencing
- Contributing members:
  - 13 FDA labs
  - 11 PulseNet labs (state public health labs)
  - 5 Dept. of Agriculture labs
  - 7 University labs
  - 1 U.S. hospital lab
  - 2 international labs (Argentina, Mexico)
  - 3 private contracting labs
- Data curation and bioinformatic support/analyses provided by National Center for Biotechnology Information (NCBI) and FDA-CFSAN.

GenomeTrakr verses PulseNet?



food/env -> GT

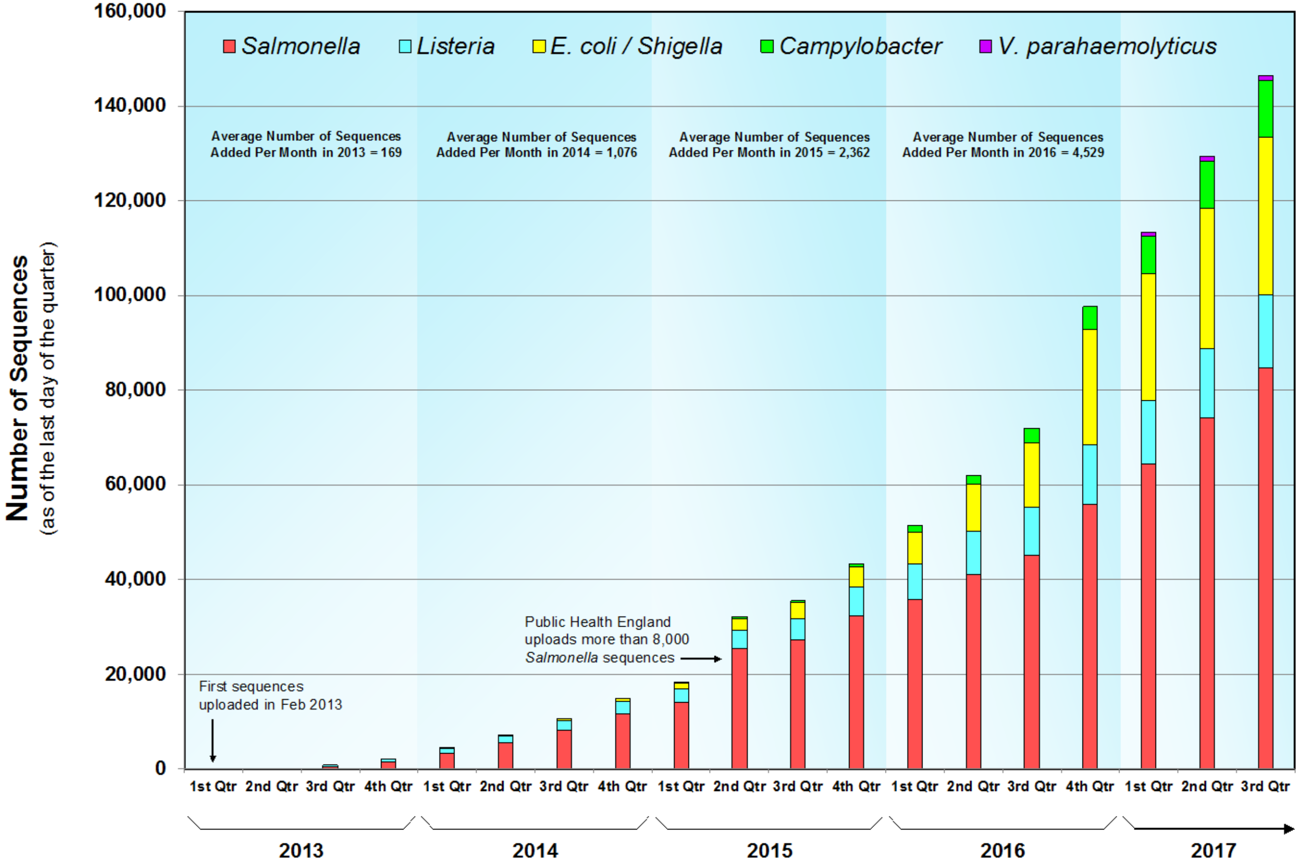
clinicals -> PN

clinicals -> PN
food/env -> GT

NCBI

Database growth:

Currently: over 150,000 genomes (all contributors)



Average Number of Sequences Added Per Month in 2013 = 169

Average Number of Sequences Added Per Month in 2014 = 1,076

Average Number of Sequences Added Per Month in 2015 = 2,362

Average Number of Sequences Added Per Month in 2016 = 4,529

Public Health England uploads more than 8,000 *Salmonella* sequences

First sequences uploaded in Feb 2013

Publicizing data

NCBI:

Sequences and metadata

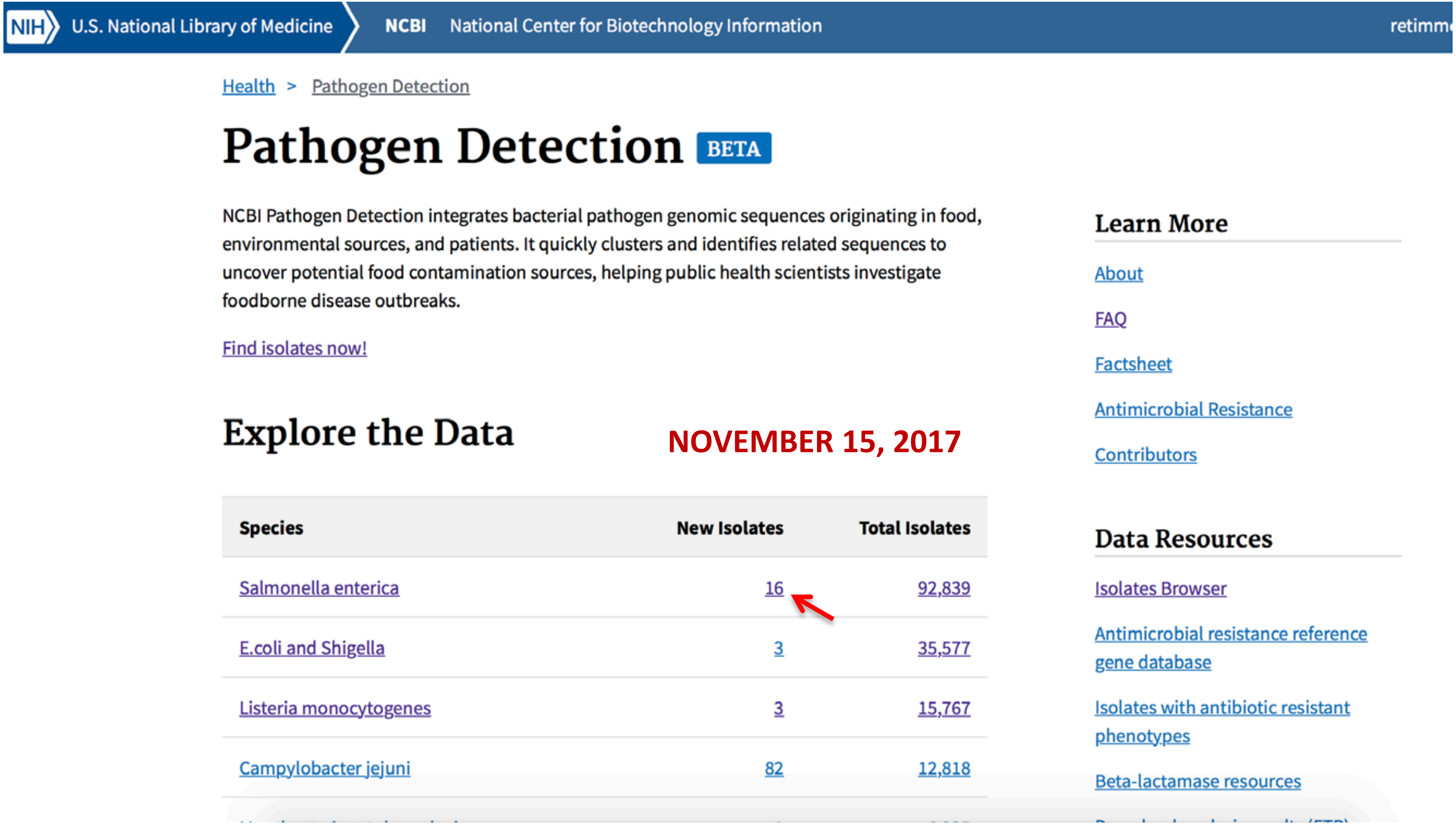
- fastq files in SRA DB, annotated assemblies in GenBank
- metadata in BioSample DB (taxonomy, collected by, country and state, year, isolation source)
- **Private**: city, county, zipcode, firm names, product names, patient data (age, sex, etc)

Analyses

- Phylogenetic trees for each pathogen published daily at NCBI: http://www.ncbi.nlm.nih.gov/projects/pathogens

GitHub:

- CFSAN SNP Pipeline: http://snp-pipeline.readthedocs.org/en/latest/index.html



NIH U.S. National Library of Medicine NCBI National Center for Biotechnology Information retimm

Health > Pathogen Detection

Pathogen Detection BETA

NCBI Pathogen Detection integrates bacterial pathogen genomic sequences originating in food, environmental sources, and patients. It quickly clusters and identifies related sequences to uncover potential food contamination sources, helping public health scientists investigate foodborne disease outbreaks.

Find isolates now!

Explore the Data

NOVEMBER 15, 2017

Species	New Isolates	Total Isolates
Salmonella enterica	16	92,839
E.coli and Shigella	3	35,577
Listeria monocytogenes	3	15,767
Campylobacter jejuni	82	12,818

Learn More

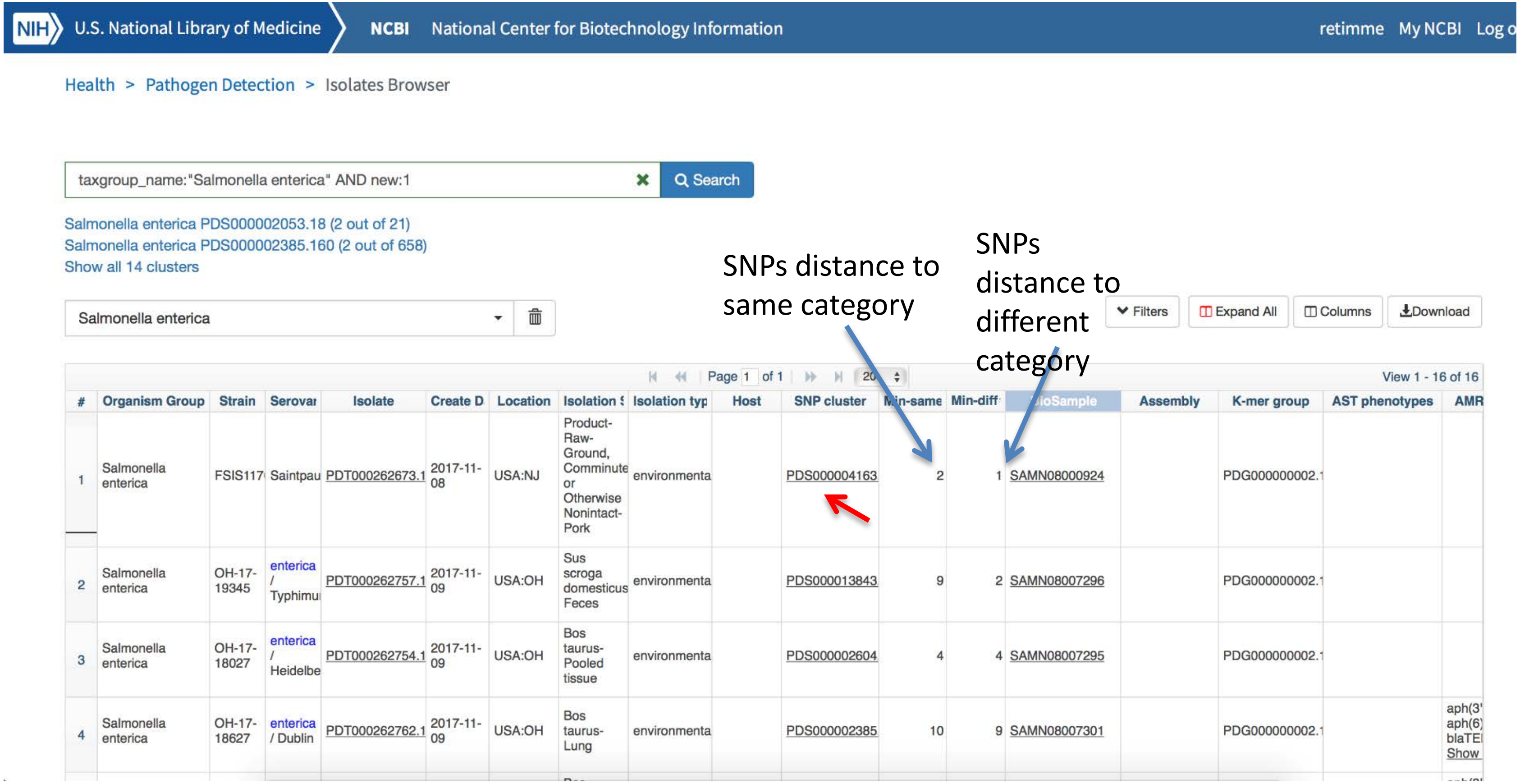
- About
- FAQ
- Factsheet
- Antimicrobial Resistance
- Contributors

Data Resources

- Isolates Browser
- Antimicrobial resistance reference gene database
- Isolates with antibiotic resistant phenotypes
- Beta-lactamase resources

http://www.ncbi.nlm.nih.gov/projects/pathogens

New isolate check - *Salmonella*



Health > Pathogen Detection > Isolates Browser

taxgroup_name:"Salmonella enterica" AND new:1

Salmonella enterica PDS000002053.18 (2 out of 21)
Salmonella enterica PDS000002385.160 (2 out of 658)
Show all 14 clusters

Salmonella enterica

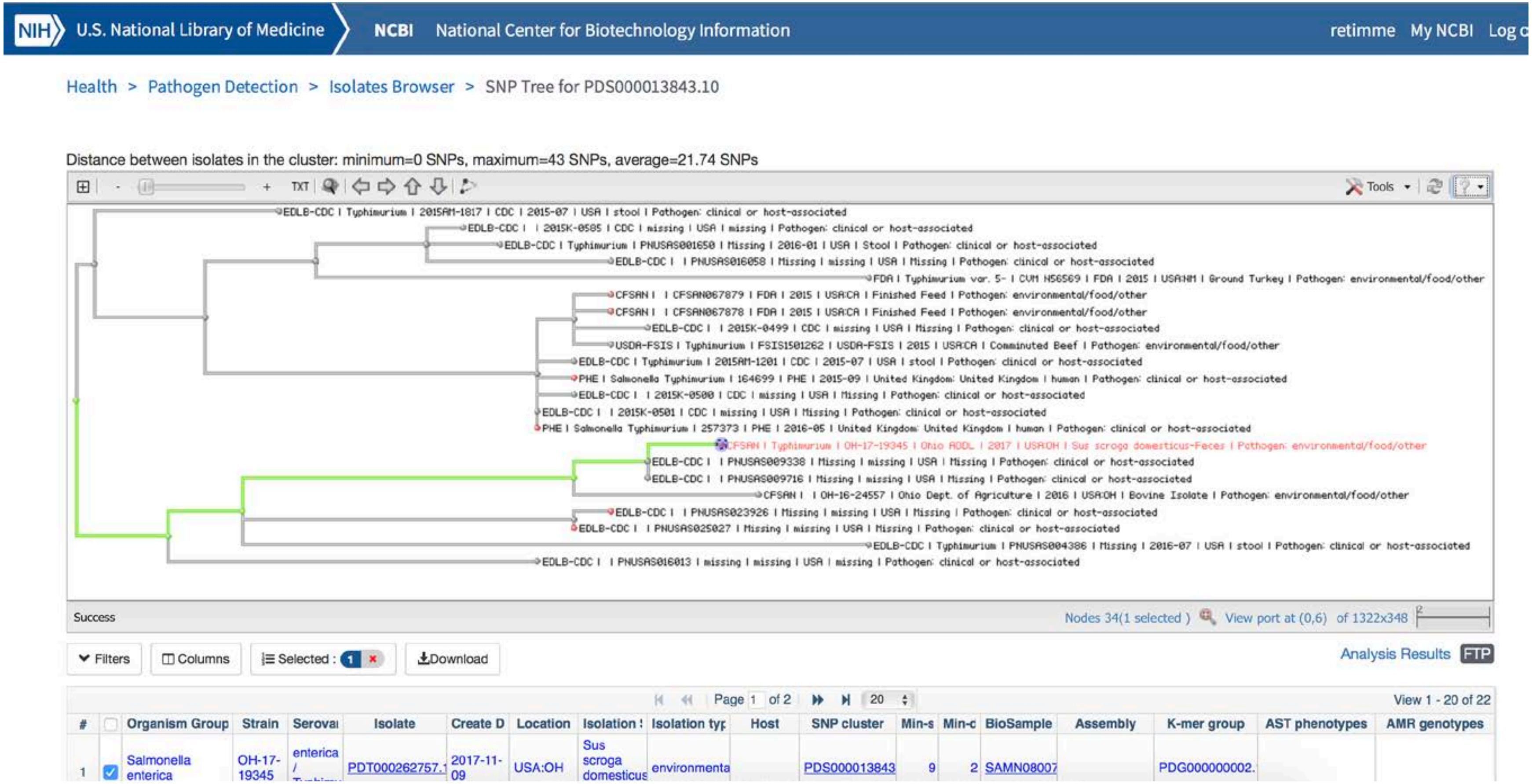
View 1 - 16 of 16

#	Organism Group	Strain	Serovar	Isolate	Create D	Location	Isolation S	Isolation typ	Host	SNP cluster	Min-same	Min-diff	BioSample	Assembly	K-mer group	AST phenotypes	AMR
1	Salmonella enterica	FSIS117	Saintpau	PDT000262673.1	2017-11-08	USA:NJ	Product-Raw-Ground, Comminute or Otherwise Nonintact-Pork	environmenta		PDS000004163	2	1	SAMN08000924		PDG000000002.1		
2	Salmonella enterica	OH-17-19345	enterica / Typhimu	PDT000262757.1	2017-11-09	USA:OH	Sus scroga domesticus Feces	environmenta		PDS000013843	9	2	SAMN08007296		PDG000000002.1		
3	Salmonella enterica	OH-17-18027	enterica / Heidelbe	PDT000262754.1	2017-11-09	USA:OH	Bos taurus-Pooled tissue	environmenta		PDS000002604	4	4	SAMN08007295		PDG000000002.1		
4	Salmonella enterica	OH-17-18627	enterica / Dublin	PDT000262762.1	2017-11-09	USA:OH	Bos taurus-Lung	environmenta		PDS000002385	10	9	SAMN08007301		PDG000000002.1		aph(3' aph(6) blaTE Show

SNPs distance to same category

SNPs distance to different category

Look at close matches within SNP cluster



Health > Pathogen Detection > Isolates Browser > SNP Tree for PDS000013843.10

Distance between isolates in the cluster: minimum=0 SNPs, maximum=43 SNPs, average=21.74 SNPs

```
EDLB-CDC | Typhimurium | 2015AM-1817 | CDC | 2015-07 | USA | stool | Pathogen: clinical or host-associated
EDLB-CDC |  | 2015K-0505 | CDC | missing | USA | missing | Pathogen: clinical or host-associated
EDLB-CDC | Typhimurium | PNUSAS001650 | Missing | 2016-01 | USA | Stool | Pathogen: clinical or host-associated
EDLB-CDC |  | PNUSAS016058 | Missing | missing | USA | Missing | Pathogen: clinical or host-associated
FDA | Typhimurium var. 5- | CVM N56569 | FDA | 2015 | USA:NM | Ground Turkey | Pathogen: environmental/food/other
CFSAN |  | CFSAN067879 | FDA | 2015 | USA:CA | Finished Feed | Pathogen: environmental/food/other
CFSAN |  | CFSAN067878 | FDA | 2015 | USA:CA | Finished Feed | Pathogen: environmental/food/other
EDLB-CDC |  | 2015K-0499 | CDC | missing | USA | Missing | Pathogen: clinical or host-associated
USDA-FSIS | Typhimurium | FSIS1501262 | USDA-FSIS | 2015 | USA:CA | Comminuted Beef | Pathogen: environmental/food/other
EDLB-CDC | Typhimurium | 2015AM-1201 | CDC | 2015-07 | USA | stool | Pathogen: clinical or host-associated
PHE | Salmonella Typhimurium | 164699 | PHE | 2015-09 | United Kingdom: United Kingdom | human | Pathogen: clinical or host-associated
EDLB-CDC |  | 2015K-0500 | CDC | missing | USA | Missing | Pathogen: clinical or host-associated
EDLB-CDC |  | 2015K-0501 | CDC | missing | USA | Missing | Pathogen: clinical or host-associated
PHE | Salmonella Typhimurium | 257373 | PHE | 2016-05 | United Kingdom: United Kingdom | human | Pathogen: clinical or host-associated
CFSAN | Typhimurium | OH-17-19345 | Ohio ADDL | 2017 | USA:OH | Sus scroga domesticus-Feces | Pathogen: environmental/food/other
EDLB-CDC |  | PNUSAS009338 | Missing | missing | USA | Missing | Pathogen: clinical or host-associated
EDLB-CDC |  | PNUSAS009716 | Missing | missing | USA | Missing | Pathogen: clinical or host-associated
CFSAN |  | OH-16-24557 | Ohio Dept. of Agriculture | 2016 | USA:OH | Bovine Isolate | Pathogen: environmental/food/other
EDLB-CDC |  | PNUSAS023926 | Missing | missing | USA | Missing | Pathogen: clinical or host-associated
EDLB-CDC |  | PNUSAS025027 | Missing | missing | USA | Missing | Pathogen: clinical or host-associated
EDLB-CDC | Typhimurium | PNUSAS004386 | Missing | 2016-07 | USA | stool | Pathogen: clinical or host-associated
EDLB-CDC |  | PNUSAS016013 | missing | missing | USA | missing | Pathogen: clinical or host-associated
```

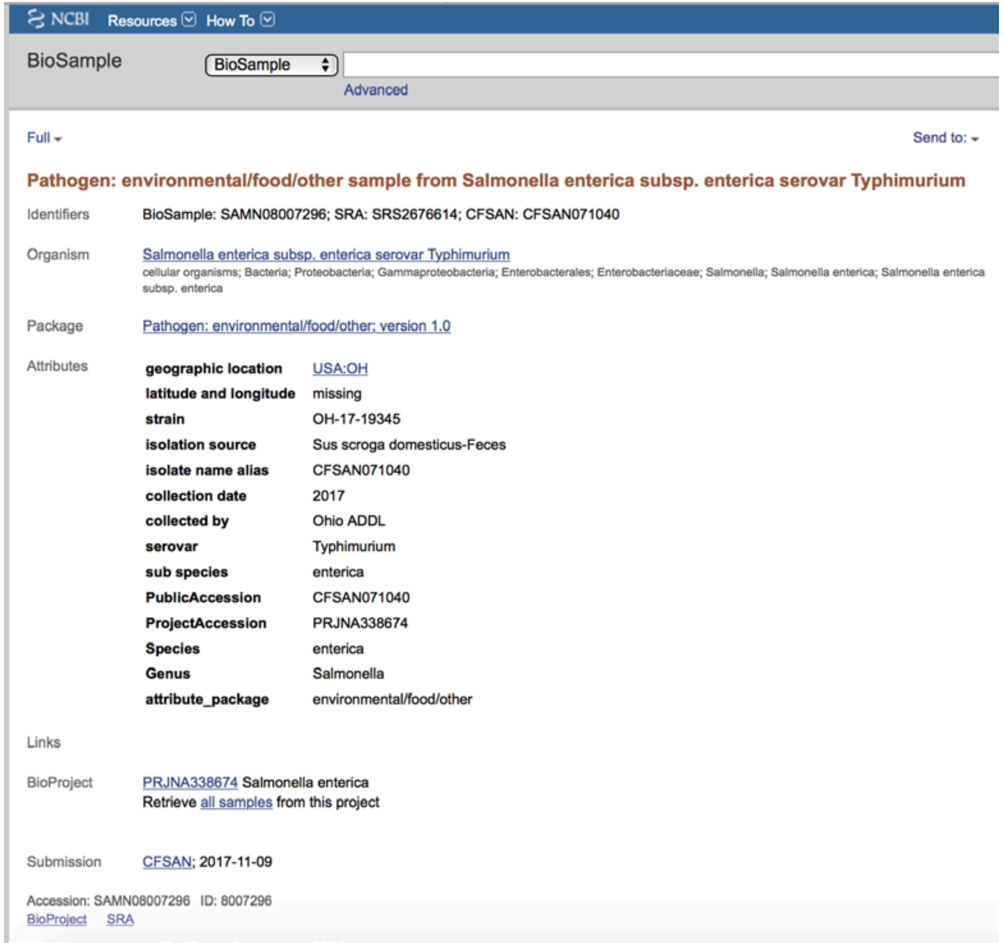
Success

Nodes 34(1 selected) View port at (0,6) of 1322x348

View 1 - 20 of 22

#	☐	Organism Group	Strain	Serovai	Isolate	Create D	Location	Isolation s	Isolation typ	Host	SNP cluster	Min-s	Min-c	BioSample	Assembly	K-mer group	AST phenotypes	AMR genotypes
1	☑	Salmonella enterica	OH-17-19345	enterica /	PDT000262757.1	2017-11-09	USA:OH	Sus scroga domesticus	environmenta		PDS000013843	9	2	SAMN08007		PDG000000002.		

Biosample: Isolate metadata



Pathogen: environmental/food/other sample from Salmonella enterica subsp. enterica serovar Typhimurium

Identifiers BioSample: SAMN08007296; SRA: SRS2676614; CFSAN: CFSAN071040

Organism Salmonella enterica subsp. enterica serovar Typhimurium
cellular organisms; Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacterales; Enterobacteriaceae; Salmonella; Salmonella enterica; Salmonella enterica subsp. enterica

Package Pathogen: environmental/food/other; version 1.0

Attributes

geographic location	USA:OH
latitude and longitude	missing
strain	OH-17-19345
isolation source	Sus scroga domesticus-Feces
isolate name alias	CFSAN071040
collection date	2017
collected by	Ohio ADDL
serovar	Typhimurium
sub species	enterica
PublicAccession	CFSAN071040
ProjectAccession	PRJNA338674
Species	enterica
Genus	Salmonella
attribute_package	environmental/food/other

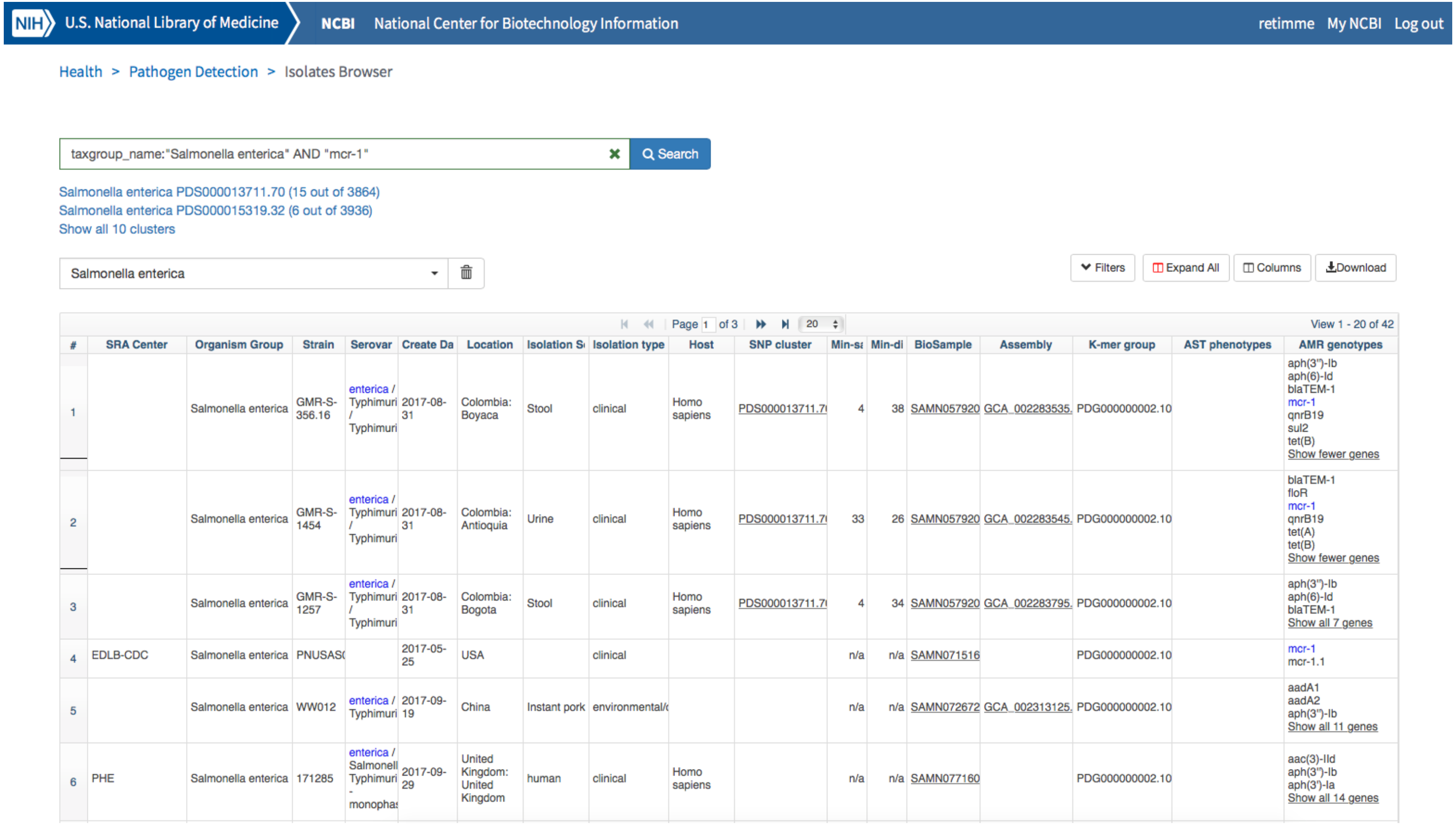
Links

BioProject PRJNA338674 Salmonella enterica
Retrieve all samples from this project

Submission CFSAN; 2017-11-09

Accession: SAMN08007296 ID: 8007296
BioProject SRA

AMR genotype prediction



Health > Pathogen Detection > Isolates Browser

taxgroup_name:"Salmonella enterica" AND "mcr-1" ✖ Search

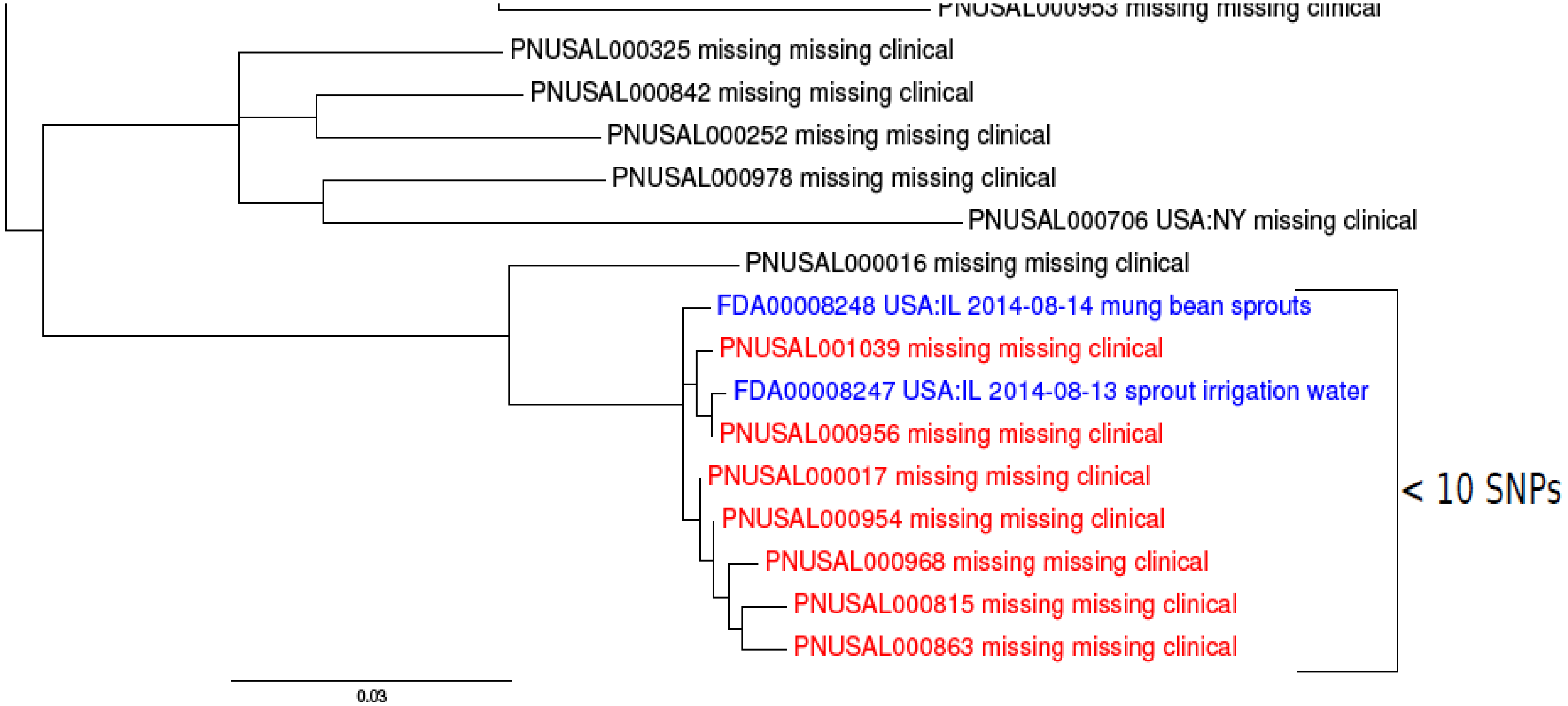
Salmonella enterica PDS000013711.70 (15 out of 3864)
Salmonella enterica PDS000015319.32 (6 out of 3936)
Show all 10 clusters

View 1 - 20 of 42

#	SRA Center	Organism Group	Strain	Serovar	Create Da	Location	Isolation S	Isolation type	Host	SNP cluster	Min-sa	Min-di	BioSample	Assembly	K-mer group	AST phenotypes	AMR genotypes
1		Salmonella enterica	GMR-S-356.16	enterica / Typhimuri / Typhimuri	2017-08-31	Colombia: Boyaca	Stool	clinical	Homo sapiens	PDS000013711.7	4	38	SAMN057920	GCA_002283535.	PDG000000002.10		aph(3")-Ib aph(6)-Id blaTEM-1 mcr-1 qnrB19 sul2 tet(B) Show fewer genes
2		Salmonella enterica	GMR-S-1454	enterica / Typhimuri / Typhimuri	2017-08-31	Colombia: Antioquia	Urine	clinical	Homo sapiens	PDS000013711.7	33	26	SAMN057920	GCA_002283545.	PDG000000002.10		blaTEM-1 floR mcr-1 qnrB19 tet(A) tet(B) Show fewer genes
3		Salmonella enterica	GMR-S-1257	enterica / Typhimuri / Typhimuri	2017-08-31	Colombia: Bogota	Stool	clinical	Homo sapiens	PDS000013711.7	4	34	SAMN057920	GCA_002283795.	PDG000000002.10		aph(3")-Ib aph(6)-Id blaTEM-1 Show all 7 genes
4	EDLB-CDC	Salmonella enterica	PNUSAS		2017-05-25	USA		clinical			n/a	n/a	SAMN071516		PDG000000002.10		mcr-1 mcr-1.1
5		Salmonella enterica	WW012	enterica / Typhimuri	2017-09-19	China	Instant pork	environmental/			n/a	n/a	SAMN072672	GCA_002313125.	PDG000000002.10		aadA1 aadA2 aph(3")-Ib Show all 11 genes
6	PHE	Salmonella enterica	171285	enterica / Salmonell Typhimuri - monophas	2017-09-29	United Kingdom: United Kingdom	human	clinical	Homo sapiens		n/a	n/a	SAMN077160		PDG000000002.10		aac(3)-IId aph(3")-Ib aph(3')-Ia Show all 14 genes

How do we use the GenomeTrakr information?

- Identify SNP cluster of interest from NCBI
- Download raw data AND run CFSAN SNP pipeline inhouse



What happens with a WGS link between a clinical and environmental sample?

- **Likely result in the following steps:**

  (1) facility/farm inspection and sampling

  (2) Pathogen positive samples are sequenced and submitted to the database

  (3) traceback/trace forward of raw materials and finished product

  (4) WGS is powerful tool that <u>supports</u> investigation

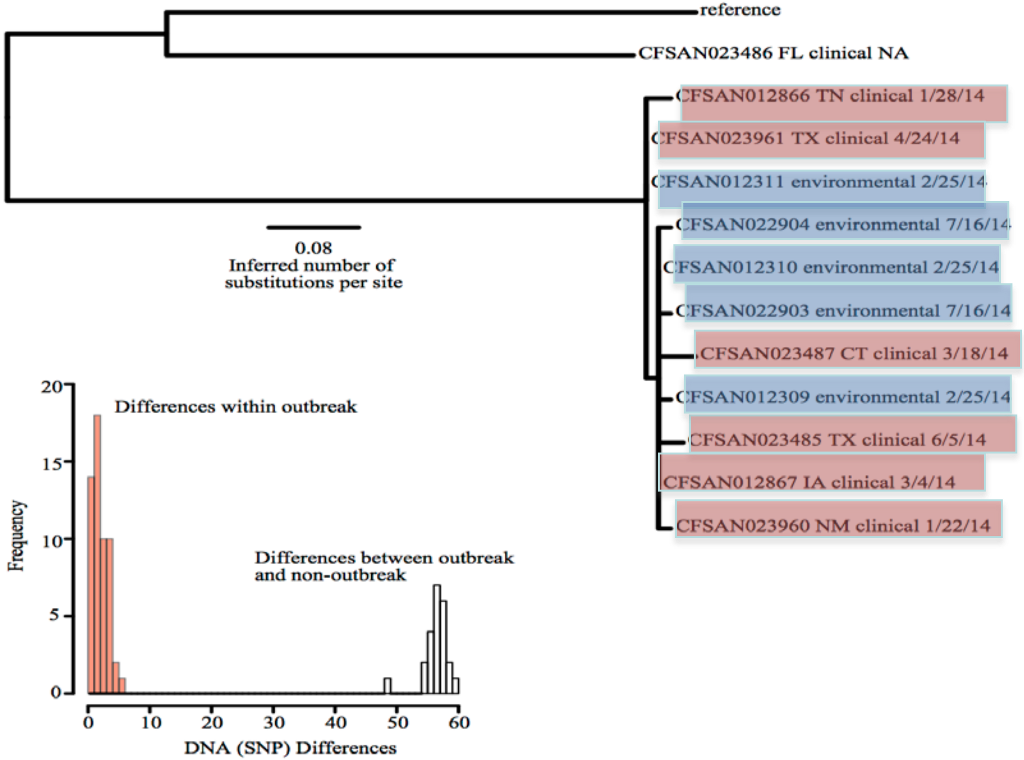
Salmonella Braenderup 2014 pre-outbreak

- In 2014, FDA conducted baseline environmental sampling in nut butter processing facilities

- A few of the samples tested positive for *S*. Braenderup and a PFGE pattern matched several cases of recent salmonellosis without a common link

- WGS was performed on both environmental and clinical isolates and found to be extremely close (2 SNP differences)

Salmonella Braenderup



🟦 env. swab

🟥 clinical

Comparing Traditional and Retrospective Outbreaks in Nut Butters

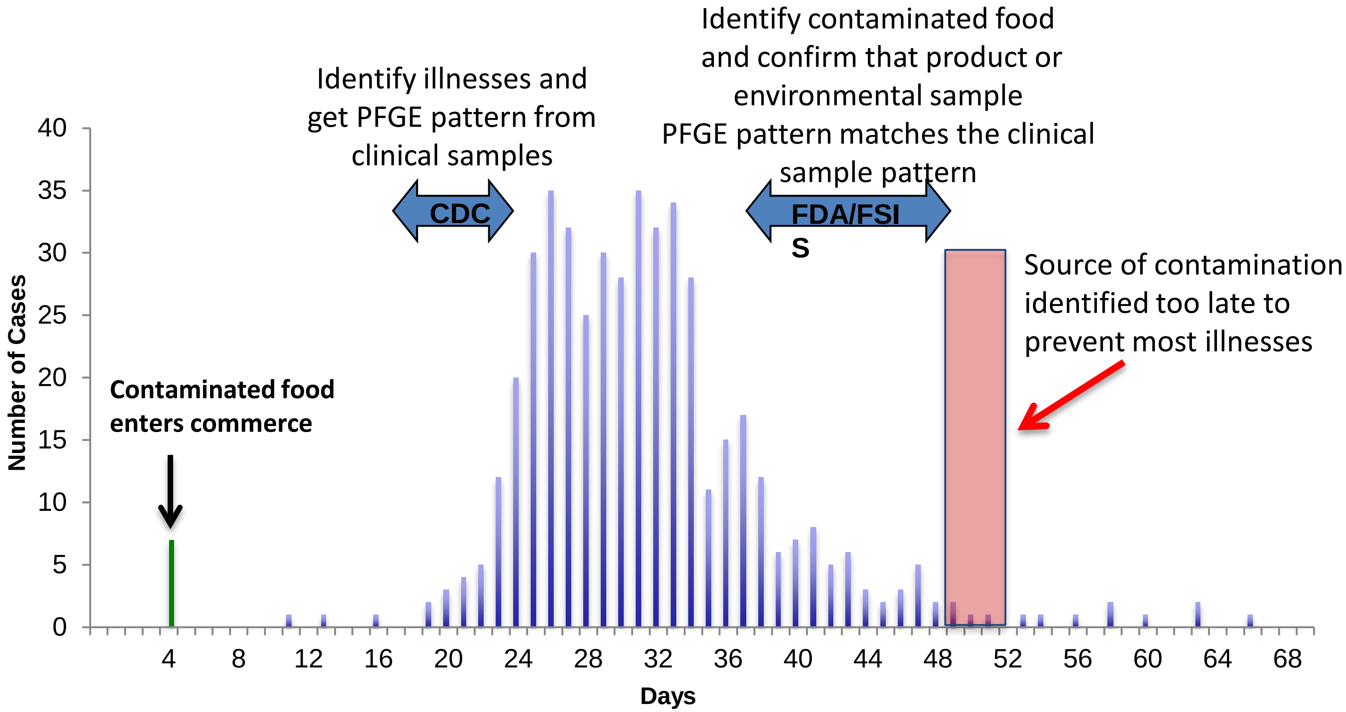
Traditional Outbreak Investigations

- *Salmonella* Tennessee (Company A, Brand A Peanut Butter, 2006/2007): 715 cases, 129 hospitalizations, no deaths
- *Salmonella* Typhimurium (Company B, Brand B Peanut butter, 2008/2009): 714 cases, 166 hospitalizations, 9 deaths
- *Salmonella* Bredeney (Company C/Brand C Peanut butter, 2012): 42 cases, 10 hospitalizations, 0 deaths

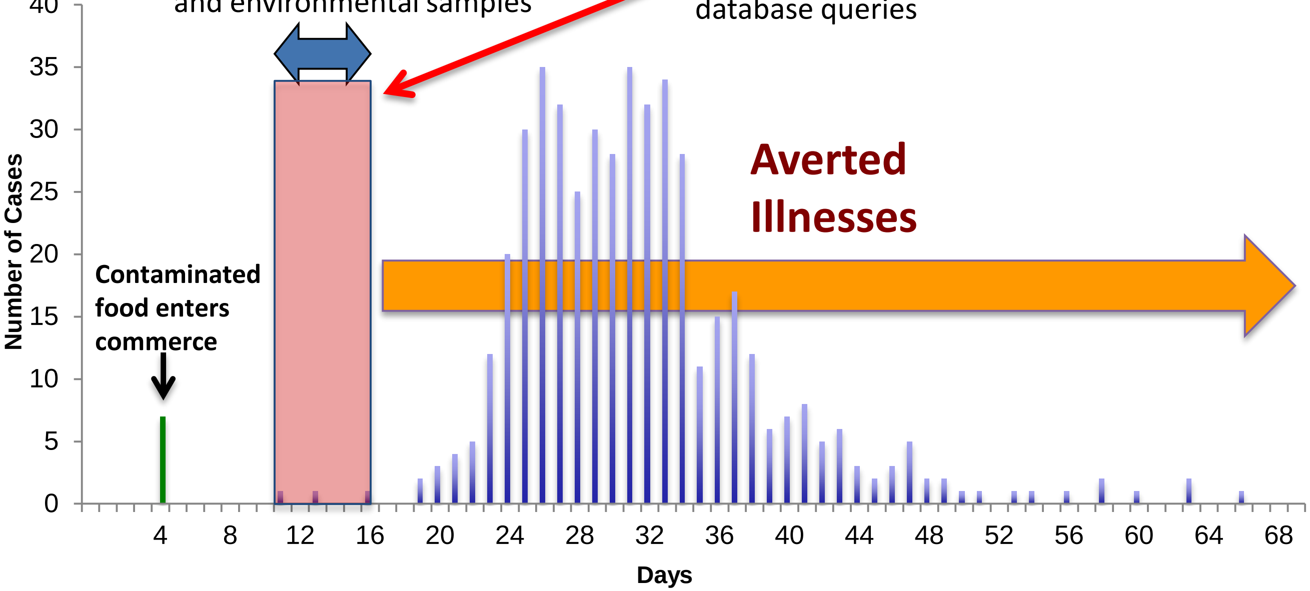
Retrospective Outbreak Investigation

- *Salmonella* Braenderup (Company D/Brand D nut butter, 2014): 6 cases, 1 hospitalization, no deaths

Timeline for Traditional Approach to Foodborne Illness Investigation



Timeline for Foodborne Illness Investigation Using Whole Genome Sequencing



FDA, CDC, FSIS, and States use WGS in real-time and in parallel on clinical, food, and environmental samples

Source of contamination identified early through WGS combined database queries

Averted Illnesses

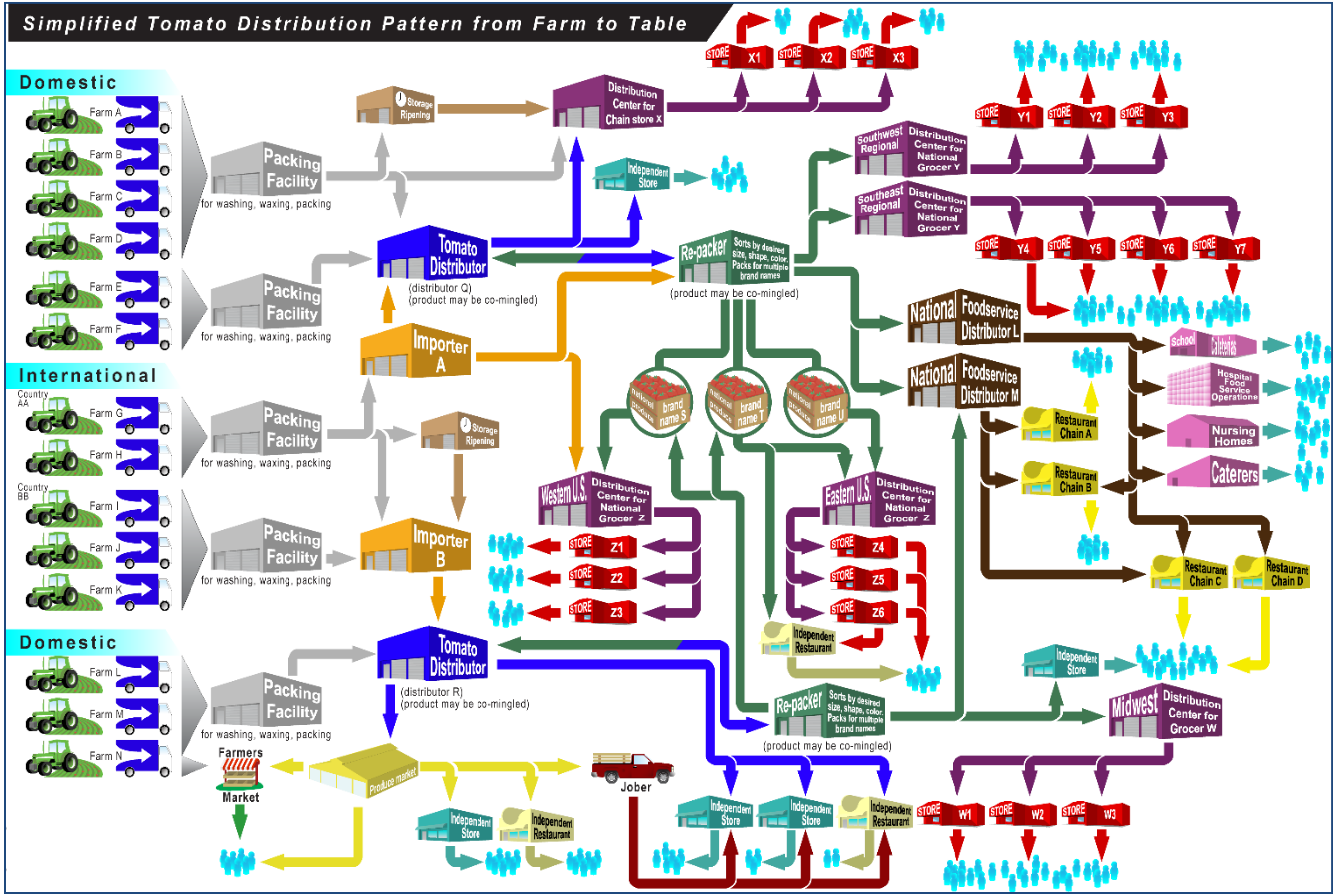
Contaminated food enters commerce

Immediate benefits of WGS to industry, growers, and distributers

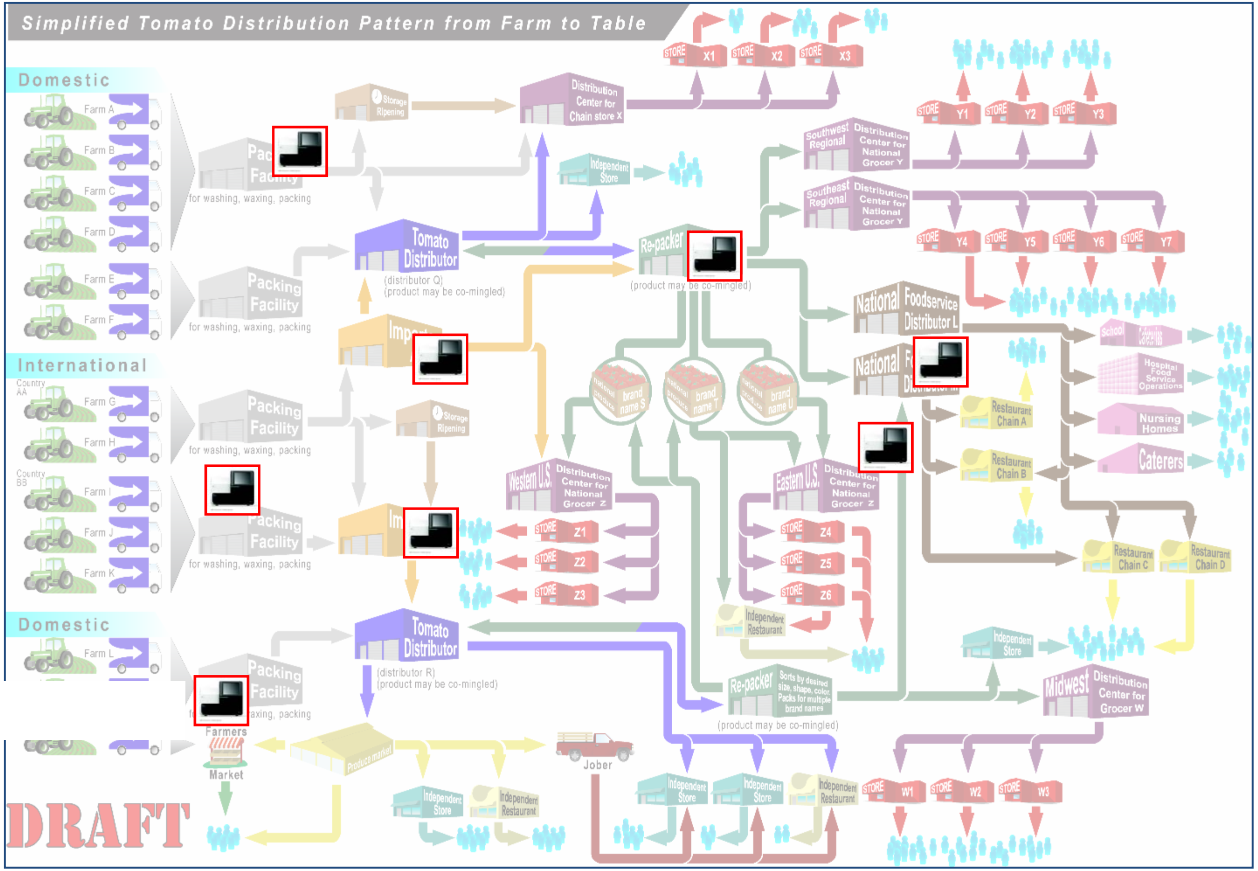
- **Earlier intervention means:**

1) Reduced amount of recalled product;
2) fewer sick patients
3) less impact overall and minimal damage to brand recognition.

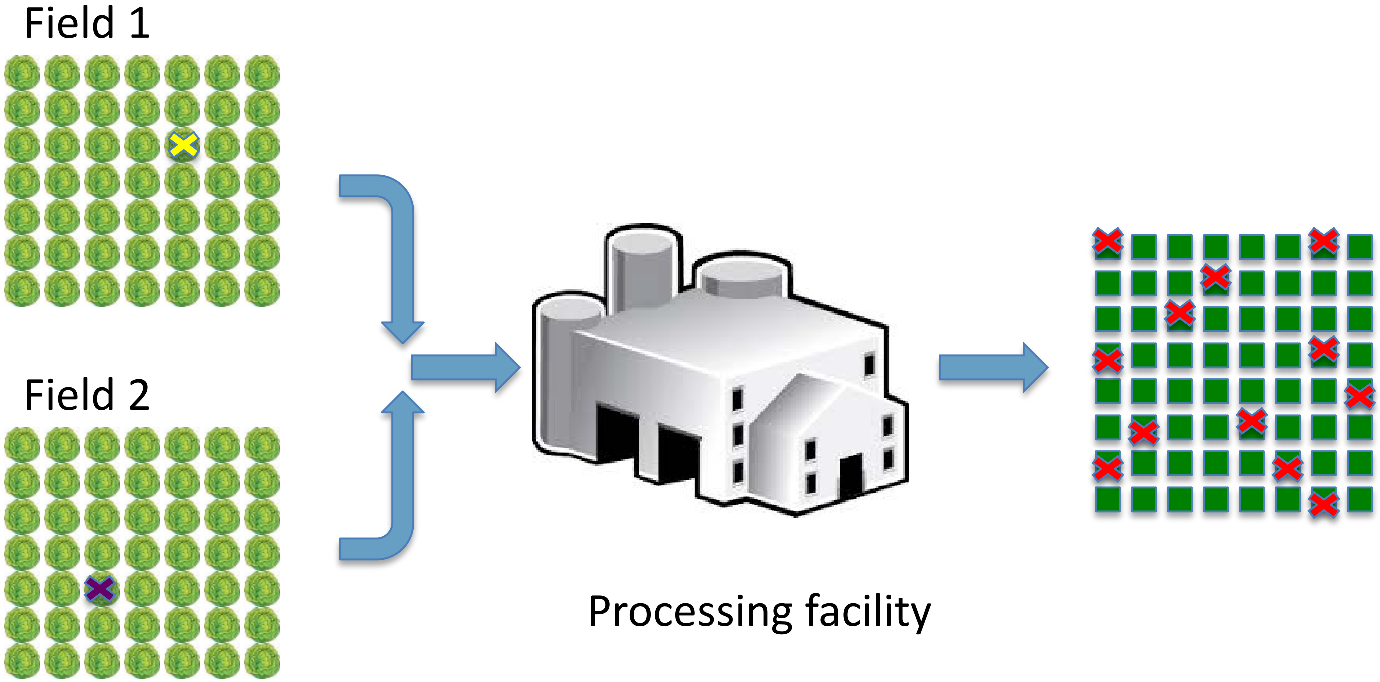
The Fresh-cut Tomato Supply Chain is complex



WGS-based monitoring can pinpoint root causes



Example 1

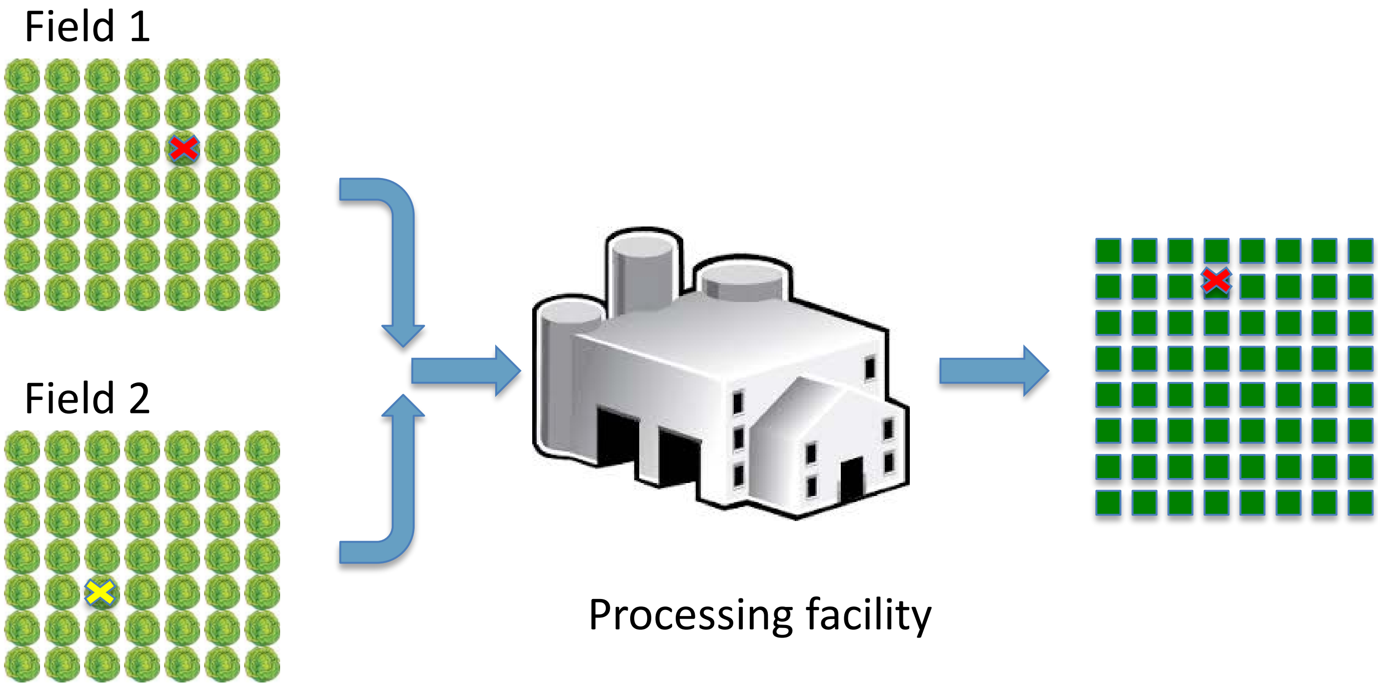


Field 1

Field 2

Processing facility

Example 2



Field 1

Field 2

Processing facility

Benefits to industry, growers, and distributers (continued)

- **Regular testing throughout network:**
  1) identifies specific suppliers that are introducing contaminants;
  2) identifies whether contaminant is resident to a facility or transient;
  3) knowledge of where contaminant is coming from allows industry to fix the problem based on scientific evidence.

  - Shift costs to the supplier who has introduced the contaminant.
  - How often is the root cause of the problem left unresolved to occur again at a later date?

One Data Record - Many Possibilities

Virulence

wgMLST

SNP

Markers

Serotype

…..AAGCTTGGAGATCTACGTGTACCTAGTCGAAGACTGAGGCTCTA….

Biofilm persistence

Resistance (Disinfectant, Heat, Heavy metal…)

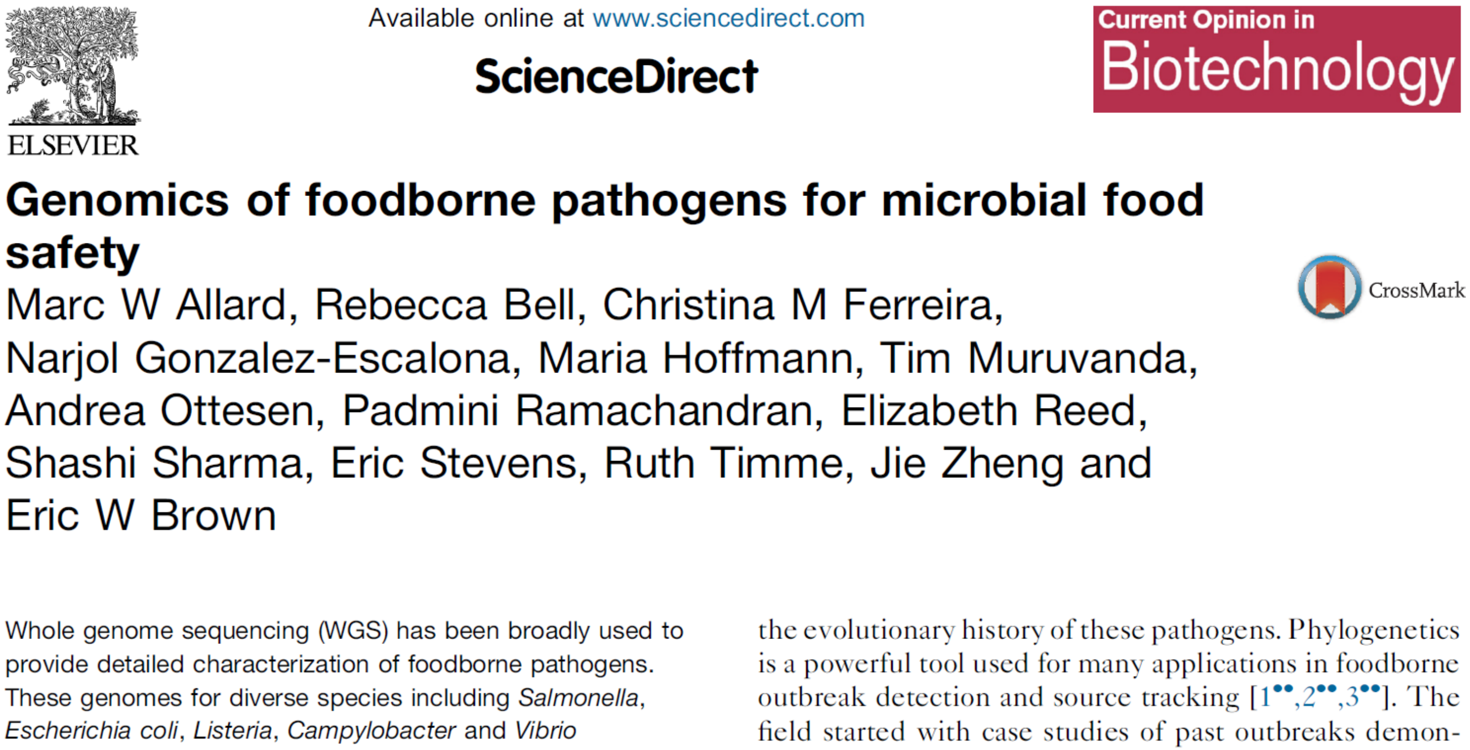
Ecological Fitness

Unknown

Adaptation

Improving Food Safety

1. Identify source of foodborne outbreaks more quickly
   - ~ WGS provides an integrated food safety surveillance system
   - ~ permits international capacity building through integration of foreign food safety entities into the GT network
2. Transparency of open data gives industry full access
   - ~ Genome data made public in real-time
   - ~ Public software and analysis tools readily available to industry for viewing of results
3. Food Safety Modernization Act (2011) – preventive Controls, Improve Industry Practices
   - ~ WGS compliments rapid testing methods with environmental monitoring for repeat positives and problems w/ resident pathogens.



Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in
Biotechnology

Genomics of foodborne pathogens for microbial food safety

Marc W Allard, Rebecca Bell, Christina M Ferreira, Narjol Gonzalez-Escalona, Maria Hoffmann, Tim Muruvanda, Andrea Ottesen, Padmini Ramachandran, Elizabeth Reed, Shashi Sharma, Eric Stevens, Ruth Timme, Jie Zheng and Eric W Brown

Whole genome sequencing (WGS) has been broadly used to provide detailed characterization of foodborne pathogens. These genomes for diverse species including *Salmonella*, *Escherichia coli*, *Listeria*, *Campylobacter* and *Vibrio*

the evolutionary history of these pathogens. Phylogenetics is a powerful tool used for many applications in foodborne outbreak detection and source tracking [1••,2••,3••]. The field started with case studies of past outbreaks demon-

Acknowledgements

- **FDA**
  - Center for Food Safety and Applied Nutrition
  - Center for Veterinary Medicine
  - Office of Regulatory Affairs

- **State Health and University Labs**
  - Alaska
  - Arizona
  - California
  - Florida
  - Hawaii
  - Maryland
  - Michigan
  - Minnesota
  - New Mexico
  - New York
  - North Carolina
  - Ohio
  - Penn State
  - South Dakota
  - Texas
  - Virginia
  - Washington

- **National Institutes of Health**
  - National Center for Biotechnology Information

- **USDA/FSIS**
  - Eastern Laboratory

- **CDC**
  - Enteric Diseases Laboratory

CFSAN contributors:

Eric Brown
Ruth Timme
Errol Strain
Marc Allard
James Pettengill
Yan Yao
Maria Sanchez-Leon

FDA U.S. FOOD & DRUG
ADMINISTRATION

Thank you!

california almonds
Almond Board of California

growing ADVANTAGE
The Almond Conference

Use **#AlmondConf** to be part of the conversation on Facebook and Twitter

What’s Next

Tuesday, December 5 at 4:15 p.m.

- State of the Industry – Hall C

Be sure to join us at 5:30 p.m. in Hall A+B for Dedicate Trade Show Time and Opening Reception, sponsored by The Bank of Stockton

CELEBRATING
1867 150 2017
YEARS
BANK OF STOCKTON